AD-A226 275

STUDIES OF GENETIC VARIATION IN THE AIDS VIRUS: RELEVANCE TO DISEASE PATHOGENESIS, ANTI-VIRAL THERAPY, AND VACCINE DEVELOPMENT

ANNUAL REPORT

GEORGE M. SHAW

JUNE 15, 1990

Supported by

90 05 035

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION	1b. RESTRICTIVE MARKINGS
Unclassified	

2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT
	approved for public release;
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	distribution unlimited

4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)

6a. NAME OF PERFORMING ORGANIZATION	6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION
University of Alabama at Birmingham		

6c. ADDRESS (City, State, and ZIP Code)	7b. ADDRESS (City, State, and ZIP Code)
701 19th Street South, LHRB 613 Birmingham, AL 35294	

8a. NAME OF FUNDING/SPONSORING ORGANIZATION U.S. Army Medical Research & Development Command	8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
		DAMD17-87-C-7038

8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS			
Fort Detrick Frederick, Maryland 21701-5012	PROGRAM ELEMENT NO.	PROJECT NO. 3M2	TASK NO.	WORK UNIT ACCESSION NO.
	623105	623105H29	AD	045

11. TITLE (Include Security Classification)
(U) Studies of Genetic Variation in the AIDS Virus: Relevance to Disease Pathogenesis, Anti-Viral Therapy, and Vaccine Development

12. PERSONAL AUTHOR(S)
George M. Shaw

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Annual	FROM 12/15/88 TO 12/14/89	1990 June 15	

16. SUPPLEMENTARY NOTATION

17.	COSATI CODES		18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	HIV, AIDS, Virology, SIV, PCR, RA 1
06	03		
06	13		

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Molecular analyses of HIV-1, HIV-2, and SIV obtained by virus culture, by lambda phage cloning from uncultured patient tissues, and by PCR (polymerase chain reaction) amplification promise to yield important and novel insights into the pathogenesis of these viruses. In year 03, work was undertaken to sequence and characterize a full-length transfection competent clone of HIV-2 (JSP4-27) representing the first *bona fide* attenuated strain of HIV-2 from Senegal, West Africa, to develop generic (universal) oligonucleotide primer pairs and "nested" PCR amplification techniques for the identification of novel strains of HIV-2 and SIV in underlined blood mononuclear cells, and to clone and characterize by lambda and PCR approaches HIV-1 from uncultured human brain. The results of these studies reported herein provide the first molecular characterization of an attenuated HIV-2 strain lacking syncytia forming properties and shows that alterations in the putative N-terminal fusion domain are not responsible for this phenotype. Secondly, the results demonstrate the successful development and use of generic PCR amplification to identify, clone, and sequence a novel subgroup of SIV_{AGM} viruses. Thirdly, full-length transfection-competent HIV-1 clones were obtained for the first time from uncultured human brain tissue and the molecular and biologic properties of these viruses were shown to be distinct in their genetic makeup and in their ability to grow in monocyte-macrophages. These latter clones of HIV-1 from human brain may be particularly important since they encode fully competent HIV-1 viruses yet have never been subjected to *in vitro* cultivation at any point prior to their cloning. In this regard, they promise to be particularly useful for studies of molecular mechanisms of viral pathogenesis *in vivo*.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION	
☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS	Unclassified	
22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE (Include Area Code)	22c. OFFICE SYMBOL
Mary Frances Bostian	301-663-7325	SGRD-RMI-S

DD Form 1473, JUN 86 — Previous editions are obsolete. — SECURITY CLASSIFICATION OF THIS PAGE

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

_____ Where copyrighted material is quoted, permission has been obtained to use such material.

_____ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

_____ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

_____ In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and Use of Laboratory Animals of the Institute of Laboratory Resources, National Research Council (NIH Publication No. 86-23, Revised 1985).

_____ For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

_____ In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institute of Health.

_____   8/29/90
PI Signature   Date

Accesion For		
NTIS CRA&I	☑	
DTIC TAB	☐	
Unannounced	☐	
Justification		
By		
Distribution /		
Availability Codes		
Dist	Avail and / or Special	
A-1		

HIV-1 and HIV-2 represent two distinct groups of human immunodeficiency viruses known to cause AIDS in infected individuals (3,11,12,21,43). While HIV-1 is the causative agent of epidemic AIDS worldwide, HIV-2 has generally been restricted to West Africa (4,7,13,28,31). Numerous isolates of HIV-1 and HIV-2 have been obtained and their biological and molecular properties characterized (1,3,12,23,43,54). Nucleotide sequence analysis shows that HIV-2 is only distantly related to HIV-1 (23), while it is more closely related to two primate retroviruses, SIV_{MAC} and SIV_{SM}, which cause an AIDS-like disease in captive macaques (8,14,27). Although genetically divergent, prototype HIV-1 and HIV-2 viruses share similar overall genomic organization and have similar biological properties, which include their propensity for rapid genetic change (18,45,54), their host cell tropism, their cytopathic effect on T-cell cultures and peripheral blood mononuclear cells *in vitro*, and their ability to form syncytia with CD4-bearing target cells (12,43). In fact, the majority of HIV-1 and HIV-2 strains isolated from patients with immunodeficiency disease have been shown to cause cell fusion and the formation of multinucleated giant cells in culture. This represents a hallmark of productive viral infection in tissue culture and accounts for the profound *in vitro* cytopathic effects of HIV (25,39,52). In contrast to HIV-1 and HIV-2 isolated from clinically ill individuals, natural infection of African green monkeys by SIV occurs with high prevalence in East and West Africa but does not lead to immunodeficiency. In addition, there is growing evidence that the biologic properties of HIV-1 *in vivo* may differ significantly from tissue culture adapted strains. Thus, in order to characterize HIVs and SIVs as they exist *in vivo*, the goals for year 03 of this contract were: (i) to characterize by genetic and biologic analysis an HIV-2 strain obtained during year 02 of this contract which showed attenuated virulence *in vitro* and *in vivo*; (ii) to develop PCR approaches that would allow, for the first time, the direct identification, cloning, and sequence analysis of SIV and HIV from uncultured blood of West African green monkeys and humans West African green monkeys and humans (all previous clones of SIV_{AGM} have been derived from East African green monkeys); (iii) to obtain by direct lambda phage

1

cloning full-length transfection competent clones of HIV-1 from uncultured human brain tissue.

Molecular Characterization of HIV-2/ST

In contrast to the prototype pathogenic HIV-1 and HIV-2 viruses, we and others have recently isolated less cytopathic strains of HIV-1 and HIV-2 which exhibit markedly different biological properties (2,9,16,34,36,50). These particular isolates cause little or no cell death in susceptible target cells, fail to induce cell fusion with CD4-bearing immortalized T-cell lines, exhibit a restricted host cell tropism with a preference for PBMC and/or macrophages, and are often derived from asymptomatic individuals. While their *in vitro* biological differences are well-documented, the genetic changes responsible for their attenuated phenotype are not understood. In order to elucidate determinants of HIV pathogenicity, we have thus begun to molecularly dissect a non-fusogenic and non-cytopathic HIV-2 isolate, termed HIV-2/ST, which was obtained from a healthy Senegalese prostitute as part of year 02 work for this contract (34). Although this virus replicated to high titers in tissue culture, it infected cells at a slower rate compared to cytopathic strains of HIV-1 and HIV-2, and caused little or no cell killing and fusion. This was the case despite the fact that its external envelope glycoprotein was cleaved correctly, transported to the cell surface, and shown to bind to a specific epitope on CD4, which was recognized by OKT4a but not OKT4 antibodies (34). HIV-2/ST therefore appeared to bind to the CD4 molecule analogous to other HIVs, but it failed to fuse with CD4-bearing target cells suggesting that its infectivity was greatly retarded at the level of cell entry (34).

Since HIV "isolates" generally represent complex mixtures of genotypically-distinct viruses, and since the biological phenotype of any HIV culture depends on the sum of the properties of each genotypic variant (22,45), we first attempted to isolate a molecular clone, which was both transfection-competent and representative of the *in vitro* properties of its parental virus. We therefore obtained three full-length proviral clones (λJSP4-27, λJSP4-32, and λJSP4-34) from a genomic library of a biologically-cloned high producer cell line, termed ST/B12, and subsequently transfected them into the neoplastic T-cell lines SupT1 (49) and

2

CEMx174 (29,47). Reverse transcriptase activity was detected in supernatants of cultures transfected with λJSP4-27 as early as five days post-transfection, while λJSP4-32 and λJSP4-34 transfected cultures revealed no signs of viral replication, indicating that these proviruses were replication-defective. Immunofluorescence analysis further confirmed the presence of virus expressing cells in λJSP4-27 transfected cultures, but failed to identify virus mediated cell fusion. Western blot analysis of purified JSP4-27 virions demonstrated a protein profile similar to that of the parental HIV-2/ST virus. To facilitate subsequent transfection experiments and to allow the direct comparison of JSP4-27 to other transfection-competent HIV-2 plasmid constructs, we sub-cloned the proviral insert of λJSP4-27 into the plasmid vector pSP65 (Figure 1A).

To test whether the transfection-derived JSP4-27 virions were infectious, filtered supernatants of plasmid-transfected cultures were transmitted to uninfected SupT1 and CEMx174 cells. The results showed that cell-free transmission of JSP4-27 virions was readily and reproducibly demonstrable. However, infection and spread in culture, particularly in SupT1 cells, occurred slowly and with considerable delay. These results were confirmed and extended in comparative studies with a transfection-derived, cytopathic HIV-2/ROD strain, termed SL1 (38). While transfection of the SL1 provirus resulted in >90% infected SupT1 or CEMx174 cultures within three to four days post-transfection, JSP4-27 transfected cultures reached only 10% infectivity in the same time period, which indicated the same reduced ability to spread in culture that had been observed for the parental HIV-2/ST virus (34). Similarly, the transfection-derived JSP4-27 cultures did not form syncytia upon cocultivation with several CD4-bearing T-cell lines, including SupT1, CEM, H9 and CEMx174 cells, whereas SL1 infected cultures produced numerous and large syncytia, as well as a profound cytopathic effect, with these same target cells (Figure 1B). These data thus showed that the JSP4-27 provirus was replication-competent and infectious, and exhibited the same non-fusogenic and non-cytopathic properties as previously described for the parental virus (34).

3

Having identified and characterized the biological features of a molecular clone of HIV-2/ST, we next sequenced its entire genome. The complete nucleotide sequence of the JSP4-27 provirus is depicted in Figure 2. The viral genome is 9,672 bp in length and exhibits an overall genomic organization of 5'LTR-gag-pol-central region-env-nef-3'LTR, which is identical to that of other cytopathic HIV-2 and SIV$_{MAC}$ proviruses. It contains all major open reading frames characteristic for HIV-2, including vpx which is present in HIV-2 and SIV$_{MAC}$ viruses but not in HIV-1 (17,26,32,53), and vpr which is present in HIV-1, HIV-2 and SIV$_{MAC}$ but not in SIV$_{AGM}$ viruses (19). Like other HIV/SIV proviruses, HIV-2/ST is flanked by LTR sequences which are known to regulate viral gene expression. Sequence comparison with other HIV-2 LTRs showed that regulatory elements, like the tata box, the polyadenylation site, core enhancer sequences, Sp1 binding sites, and the tat responsive region are all present in HIV-2/ST and that their sequences are highly conserved. The HIV-2/ST LTR is of similar length, and there are no major deletions or insertions which would distinguish it from the LTRs of other cytopathic HIV-2 viruses (data not shown).

Comparison of the deduced amino acid sequence of the HIV-2/ST reading frames suggested that they all encoded full-length and functional proteins, with the exception of the vpr gene. This open reading frame was found to contain an in-frame TAA stop codon which truncates the vpr protein prematurely after the first 32 amino acid residues. Since the JSP4-27 provirus is fully replication-competent, it can be concluded that the vpr gene product is not required for *in vitro* replication of HIV-2. This conclusion was confirmed by the biological analysis of a second vpr-deficient HIV-2 provirus independently constructed in our laboratory (J.C. Kappes and B.H. Hahn, unpublished), as well as by the findings of others (15). Moreover, since vpr-deficient proviruses of HIV-2 are also cytopathic and fusogenic, it is unlikely that the lack of a functional vpr gene in HIV-2/ST is responsible for its attenuated phenotype.

Pairwise sequence alignments of JSP4-27 to other cytopathic strains of HIV-2 similarly

4

revealed no genetic features unique for HIV-2/ST. Comparison of HIV-2/ST and HIV-2/ROD demonstrated an overall sequence divergence of 11%, which is within the expected range of genetic variability observed among geographically-distant isolates of HIV-2 (Tables 1 and 2). Three other recently reported HIV-2 viruses, HIV-2/ISY derived from a Gambian individual with AIDS (1,18), HIV-2/NIH$_z$ derived from an AIDS patient from Guinea Bissau (54), and HIV-2/GH derived from an AIDS patient from Ghana (24,30) differ from HIV-2/ROD (Cape Verde Islands, refs. 12 and 23) by 11%, 12%, and 12%, respectively. Among all these viruses, the Senegalese HIV-2/ST virus was found to be most closely related to the Gambian isolate HIV-2/ISY, which shared 90% of its nucleotide sequence with HIV-2/ST.

Since infectivity, syncytia formation, and cell fusion are viral properties which are mediated by the viral *env* gene, we examined this gene in particular with respect to sequence differences unique for HIV-2/ST. An alignment of the deduced HIV-2/ST *env* sequence to those of six other cytopathic and fusogenic HIV-2 and SIV viruses is depicted in Figure 3. Overall, the size of the various *env* sequences compared is approximately the same. In contrast to other HIV-2 and SIV$_{MAC}$ viruses, JSP4-27 contains no in-frame stop codon in its transmembrane envelope domain (Figure 3). This is consistent with the presence of a 43kD rather than a 32kD transmembrane glycoprotein on Western blots of JSP4-27 derived virions (Figure 1B), and is distinct from the protein profile of the HIV-2/ST parental strain which comprises a mixture of viruses with both full-length as well as truncated transmembrane proteins (34). Pairwise sequence alignment shows that the HIV-2/ST *env* sequence differs from other HIV-2 and SIV envelopes to the same degree as they differ from each other, with amino acid sequences varying between 16% and 30% (Table 1). 25 of 32 cysteine residues, including 22 of 23 located in the extracellular domain, are conserved among all viruses, which indicate a highly conserved envelope structure. In addition, HIV-2/ST contains 28 potential N-linked glycosylation sites which are arranged in a pattern similar to that of other viruses, and which also include one highly conserved glycosylation site previously shown to be critical for

5

HIV-1 infectivity (51). Finally, the HIV-2/ST *env* gene contains highly variable regions which correspond closely in distribution and size to similar hypervariable regions in the other *env* gene sequences.

Although a three-dimensional structure has not been determined for any HIV or SIV envelope glycoproteins, there are certain envelope domains whose functions have been characterized by mutagenesis analysis. These include the putative CD4 binding domain (35,37), the envelope precursor cleavage site (40), and the viral fusion sequence (5,6,20,35). Since sequence changes in any one of these domains could alter the fusogenic properties of a virus (35), we analyzed the envelope sequence of HIV-2/ST for particular mutations in these areas. No changes, or only conservative amino acid changes were found in an envelope domain of JSP4-27 that corresponds to the HIV-1 envelope region previously identified to be involved in CD4 binding (37). HIV-2/ST also contained an apparently functional primary envelope precursor cleavage site, with a recognition sequence (RNKR) identical to that of three other fusogenic HIV-2 or SIV viruses (compare Figure 3). In contrast to these viruses, however, HIV-2/ST was found to differ in 2 of 16 highly conserved amino acid residues at the N-terminus of the transmembrane envelope glycoprotein which, as shown by site-directed mutagenesis, contains the viral fusion domain (5). The mutations include an alanine to threonine change involving amino acid residue 517 (position 12 after the cleavage site), and a serine to alanine change involving amino acid residue 521 (position 16 after the cleavage site). Only one other fusogenic HIV-2 virus, HIV-2/ISY, contained these same changes. However, this virus exhibited three additional mutations in this same envelope area (compare Figure 3). Since the fusion domain is generally highly conserved among cytopathic HIV and SIV viruses (5), we considered the possibility that the non-fusogenic properties of HIV-2/ST resulted from these mutations.

To determine if the observed amino acid substitutions in the HIV-2/ST envelope fusion region were likely responsible for the impaired cytopathic properties of this virus, we examined

two fusogenic variants of HIV-2/ST, termed ST/24.1C and ST/24.2C (Figure 4). Both fusogenic strains were originally derived from a biologically-cloned sub-culture of HIV-2/ST, termed ST/24, which produced non-cytopathic and non-fusogenic virions biologically indistinguishable from the parental HIV-2/ST isolate (34). Following serial cell-free transmissions of ST/24 supernatant to uninfected SupT1 cells, large and numerous syncytia were observed on two independent occasions, which indicated the emergence of fusogenic progeny virus in the culture. Two cell lines were subsequently established (ST/24.1C and ST/24.2C) and shown to produce virions with fusogenic and cytopathic properties similar to prototype HIV-1 and HIV-2 viruses (J.A. Hoxie, et al., in preparation). Moreover, these cell lines were confirmed to be infected with HIV-2/ST by Southern blot analysis, which revealed no changes in their BamHI, NheI, HindIII, and PstI cleavage patterns as compared to ST/24. In order to identify the molecular basis for the phenotypical change in these variants and to determine whether a direct mutation of the viral fusion sequence had occurred, we amplified the envelope fusion domain of these cultures using the polymerase chain reaction (46).

Two oligonucleotide primers (30 mers) were designed to allow the amplification of a 544 bp envelope fragment from virally-infected cellular DNA, which included the putative precursor cleavage site as well as the envelope fusion region (Figures 5 and 6). Both primers were synthesized according to the JSP4-27 sequence, however, sequence changes were introduced so as to accommodate a BamHI site in the 5' amplimer and a Pst I site in the 3' amplimer (primer 1: 5'AGAATTGGG<u>GGATCC</u>TAAATTGATAGAAGT 3'; and primer 2: 5'GCTATTTAATTT<u>CTGCAG</u>TTCATACATGTT 3'). Total genomic DNA of ST/24, ST/24.1C, ST/24.2C, as well as ST/B12 as a control, was amplified using these primers. 100 µl of reaction mixture contained 10 mM Tris-HCl, pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 0.01% gelatine, 200 mM of each deoxynucleotide triphosphate (dNTP), 10 pM of each primer, 2.5 U of Taq polymerase and 1 µg of high molecular weight DNA. Samples were subjected to 45 amplification cycles, consisting of a denaturing step at 94°C for 90 seconds, a primer-annealing step at 50°C for 90 seconds, and a primer extension

7

step at 72°C for 135 seconds. Amplified envelope fragments were purified, cleaved with BamHI and PstI, and subsequently cloned into M13. Ten individual M13 clones per amplified DNA preparation were then isolated, and each clone was sequenced in the region, which corresponded to the N-terminus of the transmembrane envelope domain. An alignment of these sequences is depicted in Figure 5.

All 10 M13 clones derived from ST/B12 contained sequences identical to JSP4-27, which indicated that the PCR amplification procedure was reliable and did not cause frequent misincorporations of nucleotides in this particular DNA template. Sequence comparison of amplified fragments from cell line ST/24 demonstrated no differences among the individual M13 clones, but revealed 4 nucleotide point mutations between these ST/24 sequences and the JSP4-27 reference sequence. In fact, all ST/24 derived viruses, including the cytopathic and fusogenic ones, exhibited these same four nucleotide sequence differences, as well as the threonine and alanine substitutions previously identified in the envelope fusion region of JSP4-27. The results thus confirmed that the observed fusion sequence mutations were representative of all HIV-2/ST strains, regardless of their origin and biological phenotype. Interestingly, eight of nine M13 clones representing ST/24.1C and five of ten clones representing ST/24.2C contained additional point mutations, which predicted three amino acid sequence changes with respect to the ST/24 sequence (Figure 6). The presence of these mutations within the amplified material identifies the ST/24.1C and ST/24.2C viral strains as mixtures which comprise the parental ST/24 virus as well as additional genotypic variants. These newly-generated, genotypical variants must be responsible for the phenotypical changes seen in the ST/24.1C and ST/24.2C cultures. However, the genetic changes which are of biologically significance do not appear to occur in the envelope fusion domain.

While these studies were in progress, the biological properties of the HIV-2/ST envelope gene products were also analyzed in a eukaryotic expression system (42). Vaccinia virus expressed JSP4-27 envelope glycoproteins were compared to that of prototype HIV-1 and

8

HIV-2 viruses. While the processing, expression, and transport to the cell surface appeared to be unaltered, vaccinia virus-produced JSP4-27 envelope glycoproteins failed to form syncytia with CD4-bearing Hela cells. Since no other HIV-2 proteins were produced in this system, these results strongly suggested that the JSP4-27 envelope glycoproteins were primarily responsible for its non-fusogenic and non-cytopathic phenotype (42). Based on these results as well as the sequence data, we conclude that the non-fusogenic and attenuated properties of HIV-2/ST are indeed a function of its envelope gene products, although, the causative genetic defect appears not to involve mutations within the envelope fusion domain.

There are several mechanisms other than a direct alteration of the fusion sequence that could result in the biological changes observed in HIV-2/ST. These include mutations that affect envelope/CD4 interactions, mutations that cause differences in envelope glycosylation, mutations that require additional cell surface molecules to facilitate virus-cell fusion, as well as mutations that reduce the stability of envelope glycoprotein complexes on the cell surface. In fact, several naturally-occurring as well as genetically-engineered immunodeficiency viruses have been described, that are altered in their fusogenic or cytopathic properties because of one of these mechanisms. For example, possible differences in the binding affinity of HIV and SIV envelope glycoproteins to the CD4 receptor have been suggested by the finding that 25-fold more soluble CD4 is necessary to block infectivity of prototype HIV-2 viruses compared to HIV-1 (10). It is possible that the HIV-2/ST envelope glycoprotein binds the CD4 molecule with an even lower affinity, which would be expected to influence subsequent steps of viral entry, including membrane fusion and penetration. Another mechanism known to cause attenuation of virulence in naturally-occurring retroviruses involves differences in post-translational modifications of envelope glycoproteins. Mullins and coworkers showed that the pathogenic determinants of an immunodeficiency-causing FeLV-FAIDS virus were dependent on the processing of particular envelope oligosaccharides (44). Since HIV-2/ST differs in number and distribution of its potential N-linked envelope glycosylation sites from other

9

cytopathic HIV-2's, and since size differences between the exterior envelope glycoproteins of fusogenic and non-fusogenic HIV-2/ST strains have been observed (J.A. Hoxie, personal communication), a biologically-significant change in the sugar composition of the HIV-2/ST envelope cannot be excluded. Finally, a requirement of accessory molecules for virus-cell fusion represents still another potential mechanism to influence retroviral cytopathicity. Studies involving SIV_{MAC} recently revealed that this virus has a restricted host-cell range which comprises only a subset of CD4+ T-cell lines (29,33). While highly infectious and cytopathic for HUT78 and H9 cells, SIV_{MAC} does not fuse with CD4-bearing SupT1 cells. Moreover, SIV_{MAC} infects SupT1 cells only with considerable delay. It is thus conceivable, that SIV_{MAC} requires a surface molecule(s) in addition to CD4 to establish a productive infection in certain human T-cell lines. Since its infection kinetics and lack of cytopathic effect in SupT1 cells very much resemble those of HIV-2/ST, it is not unreasonable to speculate that HIV-2/ST might similarly require an additional cell surface molecule(s) for efficient cell fusion or penetration.

The availability of cytopathic variants of HIV-2/ST will be instrumental for future experiments designed to define the exact molecular determinants involved in HIV-2/ST attenuation. Molecular clones representing the fusogenic and cytopathic HIV-2/ST strains are expected to exhibit much less genetic divergence with respect to JSP4-27 than do unrelated HIV-2 viruses, like HIV-2/ROD or HIV-2/ISY. Therefore, a comparative sequence analysis is more likely to identify biologically-important differences, and the construction of chimeras between attenuated and cytopathic clones will be greatly facilitated. The fact that cytopathic and fusogenic ST/24 mutants evolved by cell free passage on two independent occasions indicates the presence of strong selective pressures for cytopathic and fusogenic viruses *in vitro*. It is possible that similar pressures are also present *in vivo* which may favor the emergence of more virulent strains in certain HIV infected individuals over time (2,9,50).

Identification and molecular characterization of a novel subgroup of SIV_{AGM} Viruses

Wild caught African green monkeys (AGM) from East and West Africa are infected with

SIV viruses at a high prevalence rate (30-60%) in contrast to Asian macaques which are not naturally infected (14,31). Because of this, it has been speculated that AGMs could harbor viruses that served as the progenitor of human immunodeficiency viruses acquired through cross-species transmission. Recent sequence analyses of a sooty mangabey virus has shown it to be most closely related to HIV-2 (27). To date, isolation and/or molecular characterization of SIV$_{AGM}$ from West Africa has not been reported. Of additional interest is the fact that wild caught AGMs, unlike experimentally infected rhesus macaques, do not appear to develop disease as a result of SIV infection. Because a molecular analysis of SIV$_{AGM}$ from West Africa could provide important information relevant to the evolutionary relationships of all HIV and SIV viruses, and because the biologic properties of these viruses are likely to be enlightening, we developed a PCR-based approach for cloning these viruses directly from uncultured AGM peripheral blood lymphocytes. Figures 7 and 8 illustrated the "nested" PCR technique that we developed which accomplishes this. Using two sequential 30 cycle rounds of amplification, first with an outer primer set and then by an inner set containing cloning sites for the M13 polylinker, we successfully cloned SIV$_{AGM}$ from four out of four seropositive AGMs and from zero of two seronegative animals. Figures 9, 10 and 11 show the nucleotide sequences of the amplified regions of the polymerase gene and their relationship to SIVs from other simian species and subspecies. These results are important for the following reasons: (i) they demonstrate for the first time the utility of the generic (universal) "nested" PCR approach for detecting and cloning novel groups of viruses whose sequence divergence from known viruses precludes conventional hybridization and cloning approaches; (ii) they identify the SIV strain infecting feral West African AGMs (sabeus) and show that this virus group is not closely related to HIV-1 or HIV-2; (iii) they provide a rapid method for obtaining molecular probes of novel viruses for full-length lambda phage cloning and other kinds of nucleic acid analyses.

Molecular Cloning and Analysis of Replication Competent HIV-1 Proviruses from Uncultured Human Brain in AIDS Dementia Complex

11

All previous full-length replication-competent proviral clones of HIV-1 have been obtained from cell cultures of amplified virus, due to the low abundance of viral DNA in chronically infected individuals. Because such virus strains may be altered be a result of *in vitro* propagation, we sought to molecularly clone and genetically and biologically characterize full-length replication competent HIV-1 proviruses directly from uncultured human brain tissue of a patient with AIDS Dementia Complex (ADC). Objectives were (i) to generate replication competent proviral clones from uncultured human tissue thereby allowing an analysis of genome organization and gene structure-function relationships of virus not subjected to *in vitro* selection pressures, (ii) to generate transfection-derived genetically defined HIV-1 strains without interim cell culture for analysis of virus biological properties, replicative DNA intermediate forms, integration status, and the possible existence of defective and/or helper forms, (iii) to compare and characterize genotypic variation of proviral clones obtained by direct lambda phage cloning as compared to clones obtained by PCR amplification. For these studies, high molecular weight DNA from a brain specimen obtained at necropsy was subjected to lambda phage cloning and 10 HIV-1 proviral clones out of 8×10^6 recombinants were obtained (Figure 12). These proviral clones contained integrated and unintegrated forms, forms with one or two LTRs, and genomes with large deletions and self-integrated, reversed LTR sequences. Eight HIV-1 proviral clones (in lambda) and 11 PCR derived clones from the same brain were sequenced in a 525 bp hypervariable envelope region. All 19 clones were highly related yet distinct with nucleotide variation between 0.1 and 0.3%. Four proviral clones in lambda were determined to be full-length by restriction mapping and two of these were shown to be transfection-competent in Cos-1 cells and replication competent in human T cells and monocytes after cell free passage. We have found in preliminary studies that these molecularly derived strains of HIV-1 from uncultured human brain replicate to higher levels in monocyte-macrophages than other putative monocyte tropic viruses and therefore they represent important reagents for characterizing monocyte tropic viral determinants at a

molecular level. The genomic organization, structure-function characteristics of specific gene products, and the biological properties of these HIV-1 proviral clones are under study.

REFERENCES

1. **Albert, J., U. Bredberg, F. Chiodi, B. Bottiger, E.M. Fenyo, E. Norrby and G. Biberfeld.** 1987. A new human retrovirus isolate of West African origin (SBL-6669) and its relationship to HTLV-IV, LAV-II, and HTLV-IIIB. AIDS Res. Human Retro. 3:3-10.

2. **Asjo, B., J. Albert, A. Karlsson, L. Morfeldt-Mamson, G. Biberfeld, K. Lidman, and E.M. Fenyo.** 1986. Replicative properties of human immunodeficiency virus from patients with varying severity of HIV infection. Lancet II:660-662.

3. **Barre-Sinoussi, F., J.C. Chermann, F. Rey, M.T. Nugeyre, S. Chamaret, J. Gruest, C. Dauguet, C. Axler-Blin, F. Vezinet-Brun, C. Rouzioux, W. Rozenbaum and L. Montagnier.** 1983. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). Science **220**:868-870.

4. **Berkelman, R.L., W.L. Heyward, J.K. Stehr-Green and J.W. Curran.** 1989. Epidemiology of human immunodeficiency virus infection and acquired immunodeficiency syndrome. Amer. J. of Med. **86**:761-770.

5. **Bosch, M.L., P.L. Earl, K. Fargnoli, S. Picciafuoco, F. Giombini, F. Wong-Staal and G. Franchini.** 1989. Identification of the fusion peptide of primate immunodeficiency viruses. Science **244**:694-697.

6. **Brasseur, R., B. Cornet, A. Burny, M. Vandenbranden and J.M. Ruysschaert.** 1988. Mode of insertion into a lipid membrane of the N-terminal HIV gp41 peptide segment. AIDS Res. and Human Retro. **4**:83-90.

7. **Centers for Disease Control:** AIDS due to HIV-2 infection -- New Jersey. MMWR 1988; **37**:33-35.

8. Chakrabarti, L., M. Guyader, M. Alizon, M.D. Daniel, R.C. Desrosiers, P. Tiollais, P. Sonigo. 1987. Sequence of simian immunodeficiency virus from macaque and its relationship to other human and simian retroviruses. Nature **328**:543-547.

9. Cheng-Mayer, C., D. Seto, M. Tateno, and J.A. Levy. 1988. Biologic features of HIV-1 that correlate with virulence in the host. Science **240**:80-82.

10. Clapham, P.R., J.N. Weber, D. Whitby, K. McIntosh, A.G. Dalgleish, P.J. Maddon, K.C. Deen, R.W. Sweet and R.A. Weiss. 1989. Soluble CD4 blocks the infectivity of diverse strains of HIV and SIV for T-Cells and monocytes but not for brain and muscle cells. Nature **337**:368-370.

11. Clavel, F., K. Mansinho, S. Chamaret, D. Guetard, V. Favier, J. Nina, M.O. Santos-Ferreira, J.L. Champalimaur and L. Montagnier. 1987. Human immunodeficiency virus type 2 infection associated with AIDS in West Africa. New England J. Med. **316**:1180-1185.

12. Clavel, F., D. Guetard, F. Brun-Vezinet, S. Chamaret, M.A. Rey, M.O. Santos-Ferreira, A.G. Laurent, C. Dauguet, C. Katlama, C. Rouzioux, D. Klatzmann, J. Champalimaud and L. Montagnier. 1986. Isolation of a new human retrovirus from West African patients with AIDS. Science **233**:343-346.

13. Cortes, E., R. Detels, D. Aboulafia, X.L. Li, T. Moudgil, M. Alam, C. Bonecker, A. Gonzaga, L. Oyafuso, M. Tondo, C. Boite, N. Hammershlak, C. Capitani, D.J. Slamon and D.D. Ho. 1989. HIV-1, HIV-2, and HTLV-1 infection in high-risk groups in Brazil. New England J. Med. **320**:953-958.

14. Daniel, M.D., N.L. Letvin, N.W. King, M. Kannagi, P.K. Sehgal, R.D. Hunt, P.J. Kanki, M. Essex and R.C. Desrosiers. 1985. Isolation of T-cell tropic HTLV-III-like retrovirus from macaques. Science **228**:1201-1204.

15. **Dedera, D., W. Hu, N.V. Heyden and L. Ratner.** 1989. Viral protein R of human immunodeficiency virus types 1 and 2 is dispensable for replication and cytopathogenicity in lymphoid cells. J. Virol. **63**:3205-3208.

16. **Evans, L.A., J. Moreau, K. Odehourl, H. Legg, A. Barboza, C. Cheng-Mayer and J.A. Levy.** 1988. Characterization of a noncytopathic HIV-2 strain with unusual effects on CD4 expression. Science **240**:1522-1525.

17. **Franchini, G., J.R. Rusche, T.J. O-Keeffe and F. Wong-Staal.** 1988. The human immunodeficiency virus type 2 (HIV-2) contains a novel gene encoding a 16 kD protein associated with mature virions. AIDS Res. Hum. Retro. **4**:243-250.

18. **Franchini, G., K.A. Fargnoll, F. Glombini, L. Jagodzinski, A. De Rossi, M. Bosch, G. Biberfeld, E.M. Fenyo, J. Albert, R.C. Gallo and F. Wong-Staal.** 1989. Molecular and biological characterization of a replication competent human immunodeficiency type 2 (HIV-2) proviral clone. Proc. Natl. Acad. Sci. USA **86**:2433-2437.

19. **Fukasawa, M., T. Miura, A. Hasegawa, S. Morikawa, H. Tsujimoto, K. Miki, T. Kitamura and M. Hayami.** 1988. Sequence of simian immunodeficiency virus from African green monkey, a new member of the HIV/SIV group. Nature **333**:457-461.

20. **Gallaher, W.R.** 1987. Detection of a fusion peptide sequence in the transmembrane protein of human immunodeficiency virus. Cell **50**:327-328.

21. **Gallo, R.C., S.Z. Salahuddin, M. Popovic, G.M. Shearer, M. Kaplan, B.F. Haynes, T.J. Palker, R. Redfield, J. Oleske, B. Safal, G. White, P. Foster and P.D. Markham.** 1984. Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. Science **224**:500-503.

22. **Goodenow, M., T. Huet, W. Saurin, S. Kwok, J. Sninsky and S. Wain-Hobson.** 1989. HIV-1 isolates are rapidly evolving quasispecies: evidence for viral mixtures and preferred nucleotide substitutions. J. AIDS **2**:344-352.

23. Guyader, M., M. Emerman, P. Sonigo, F. Clavel, L. Montagnier and M. Alizon. 1987. Genome organization and transactivation of the human immunodeficiency virus type 2. Nature **326**:662-669.

24. Hasegawa, A., H. Tsujimoto, N. Maki, K. Ishikawa, T. Miura, M. Fukasawa, K. Miki and M. Hayami. 1989. The genomic divergence of HIV-2 from Ghana. AIDS and Hum. Retro. (in press).

25. Haseltine, J.G. and J. Sodroski. 1987. Cell membrane fusion mediated by the envelope glycoprotein as the primary effector of AIDS virus cytopathicity. In: Acquired Immunodeficiency Syndrome. Eds. J.C. Gluckman and E. Vilmer. (Elsevier) pp. 47-56.

26. Henderson, L.E., R.C. Sowder, T.D. Copeland, R.E. Benvenisite and S. Oroszlan. 1988. Isolation and characterization of a novel protein (X-ORF product) from SIV and HIV-2. Science **241**:199-201.

27. Hirsch, V.M., R.A. Olmsted, M. Murphey-Corb, R.H. Purcell and P.R. Johnson. 1989. An African primate lentivirus (SIV$_{sm}$) closely related to HIV-2. Nature **339**:389-392.

28. Horsburgh, C.R., Jr. and S.D. Holmberg. 1988. The global distribution of human immunodeficiency virus type-2 (HIV-2) infection. Transfusion **28**:192-195.

29. Hoxie, J.A., B.S. Haggarty, S.E. Bonser, J.L. Rackowski, H. Shan and P.J. Kanki. 1988. Biological characterization of a simian immunodeficiency virus-like retrovirus (HTLV-IV): Evidence for CD4-associated molecules reauired for infection. J. Virol. **62**:2557-2568.

30. Ishikawa, K., H. Tsujimoto, M. Nakai, J.A.A. Mingle, M. Osei-Kwasi, S.E. Aggrey, V.B.A. Nettey, S.N. Afoakwa, M. Fukasawa, T. Kodama, M. Kawamura and M. Hayami. 1988. Isolation and characterization of HIV-2 from an AIDS patient in Ghana. AIDS **2**:383-388.

31.   **Kankl, P.J., S. M'Boup, D. Ricard, F. Brin, F. Denis, C. Boye, L. Sangare, K. Travers, M. Albaum, R. Marlink, J.L. Romet-Lemonne and M. Essex.** 1987. Human T-lymphotropic virus type 4 and the human immunodeficiency virus in West Africa. Science **236**:827-831.

32.   **Kappes, J.C., C.D. Morrow, S.W. Lee, B.A. Jameson, S.B.H. Kent, L.E. Hood, G.M. Shaw and B.H. Hahn.** 1988. Identification of a novel retroviral gene unique to human immunodeficiency virus type 2 and simian immunodeficiency virus SIV_{MAC}. J. Virol. **62**:3501-3505.

33.   **Koenig, S., V.M. Hirsch, R.A. Olmsted, D. Powell, W. Maury, A. Rabson, A.S. Fauci, R.H. Purcell and P.R. Johnson.** 1989. Selective infection of human CD4+ cells by simian immunodeficiency virus: Productive infection associated with envelope glycoprotein-induced fusion. Proc. Natl. Acad. Sci. USA **86**:2443-2447.

34.   **Kong, L.I., S.W. Lee, J.C. Kappes, J.S. Parkin, W.D. Decker, J.A. Hoxie, B.H. Hahn and G.M. Shaw.** 1988. West African HIV-2-related human retrovirus with attenuated cytopathicity. Science **240**:1525-1529.

35.   **Kowalski, M., J. Potz, L. Basipripour, T. Dorfman, W.C. Goh, E. Terwilliger, A. Dayton, C. Rosen, W. Haseltine and J. Sodroski.** 1987. Functional regions of the envelope glycoprotein of human immunodeficiency virus type 1. Science **237**:1351-1355.

36.   **Kuhnel, H., H. von Briesen, U. Dietrich, M. Adamski, D. Mix, L. Biesert, R. Kreutz, A. Immelmann, K. Henco, C. Meichsner, R. Andreesen, H. Gelderblom, H. Rubsamen-Waigmann.** 1989. Molecular cloning of two West African human immunodeficiency virus type 2 isolates that replicate well in macrophages: A Gambian isolate, from a patient with neurologic acquired immunodeficiency syndrome, and a highly divergent Ghanian isolate. Proc. Natl. Acad. Sci. USA **86**:2383-2387.

37. **Lasky, L.A., G. Nakamura, D.H. Smith, C. Fennie, C. Shimasaki, E. Patzer, P. Berman, T. Gregory and D.J. Capon.** 1987. Delineation of a region of the human immunodeficiency virus type 1 gp120 glycoprotein critical for interaction with the CD4 receptor. Cell **50**:975-985.

38. **Lee, S.W., J.C. Kappes, W.D. Decker, J. Park, C.D. Morrow, S. Su, L.I. Kong, R.C. Derosiers, G.M. Shaw and B.H. Hahn.** 1988. The role of vpx in the life cycle of HIV-2. In: Proceedings of the Colloque des Cent Gardes, Paris, France, pp. 11-17.

39. **Lifson, J.D., G.R. Reyes, M.S. McGrath, B.S. Stein and E.G. Engleman.** 1986. AIDS retrovirus induced cytopathology: Giant cell formation and involvement of CD4 antigen. Science **232**:1123-1127.

40. **McCune, J.M., L.B. Rabin, M.B. Feinberg, M. Lieberman, J.C. Kosek, G.R. Reyes and Weissman.** 1988. Endoproteolytic cleavage of gp160 is required for the activation of human immunodeficiency virus. Cell **53**:55-67.

41. **Maxam, A.M. and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499-560.

42. **Mulligan, M.J., P. Kumar, H. Hui, R.J. Owens, G.D. Ritter, Jr., B.H. Hahn and R.W. Compans.** The *env* product of an infectious noncytopathic HIV-2 is deficient in syncytium formation. (submitted).

43. **Popovic, M., M.G. Sarngadharan, E. Read and R.C. Gallo.** 1984. Detection, isolation and continuous production of cytopathic retroviruses (HTLV-III) from patients with AIDS and pre-AIDS. Science **224**:497-500.

44. **Poss, M.L., J.I. Mullins and E.A. Hoover.** 1989. Posttranslational modifications distinguish the envelope glycoprotein of the immunodeficiency disease-inducing feline leukemia virus retrovirus. J. Virol. **63**:189-195.

45. **Saag, M.S., B.H. Hahn, J. Gibbons, Y. Li, E.S. Parks, W.P. Parks and G.M. Shaw.** 1988. Extensive variation of human immunodeficiency virus type-1 *in vivo*. Nature 334:440-444.

46. **Saiki, R.D., D.H. Gelfand, S. Stoffel, S.J. Scharf, R. Higuchi, G.T. Horn, K.B. Mullis and H.A. Erlich.** 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239:487-491.

47. **Salter, R.D., D.M. Howell and P. Cresswell.** 1985. Genes regulating HLA class-I antigen expression in T-B lymphoblast hybrids. Immunogenetics 21:235-246.

48. **Sanger, F., S. Nicklen and A.R. Coulson.** 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad Sci. USA 74:5463-5467.

49. **Smith, S.D., M. Shatsky, P.S. Cohen, R. Warnke, M.P. Link and B.E. Glader.** 1984. Monoclonal antibody and enzymatic profiles of human malignant lymphoid T-cells and derived cell lines. Cancer Research 44:5657-5660.

50. **Tersmette, M., R.A. Gruters, F. de Wolf, R.E.Y. de goede, J.M.A. Lange, P.T.A. Schellekens, J.Goudsmit, H.G. Huisman and F. Miedema.** 1989. Evidence for a role of virulent human immunodeficiency virus (HIV) variants in the pathogenesis of acquired immunodeficiency syndrome: studies on sequential HIV isolates. J. Virol. 63:2118-2125.

51. **Willey, R.L., D.H. Smith, L.A. Lasky, T.S. Theodore, P.L. Earl, B. Moss, D.J. Capon and M.A. Martin.** 1988. In vitro mutagenesis identifies a region within the envelope gene of the human immunodeficiency virus that is critical for infectivity. J. Virology 62:139-147.

52. **Yoffe, B., D.E. Lewis, B.L. Petrie, C.A. Noonan, J.L. Melnick and F.B. Hollinger.** 1987. Fusion as a mediator of cytolysis in mixtures of uninfected CD4+ lymphocytes and cells infected by human immunodeficiency virus. Proc. Natl. Acad. Sci. USA 84:1429-1433.

53. **Yu, X.F., S. Ito, M. Essex and T.H. Lee.** 1988. A naturally immunogenic virion-associated protein specific for HIV-2 and SIV. Nature **335**:262-265.

54. **Zagury, J.F., G. Franchini, M. Reitz, E. Collaiti, B. Starcich, L. Hall, K. Fargnoli, L. Jagodzinski, H.G. Guo, F. Laure, S.K. Arya, S. Josephs, D. Zagury, F. Wong-Staal and R.C. Gallo.** 1988. Genetic variability between isolates of human immunodeficiency virus (HIV) type 2 is comparable to the variability among HIV type 1. Proc. Natl. Acad. Sci. USA **85**:5941-5945.

A: Construction of a replication-competent HIV-2/ST plasmid clone

B: Syncytia Formation Assay

HIV-2/ST
pJSP4-27

HIV-2/ROD
pSL1

C: Western Blot

JSP SL1

gp43 —
—gp32
p24 —
—p24

human
anti-HIV-2 serum
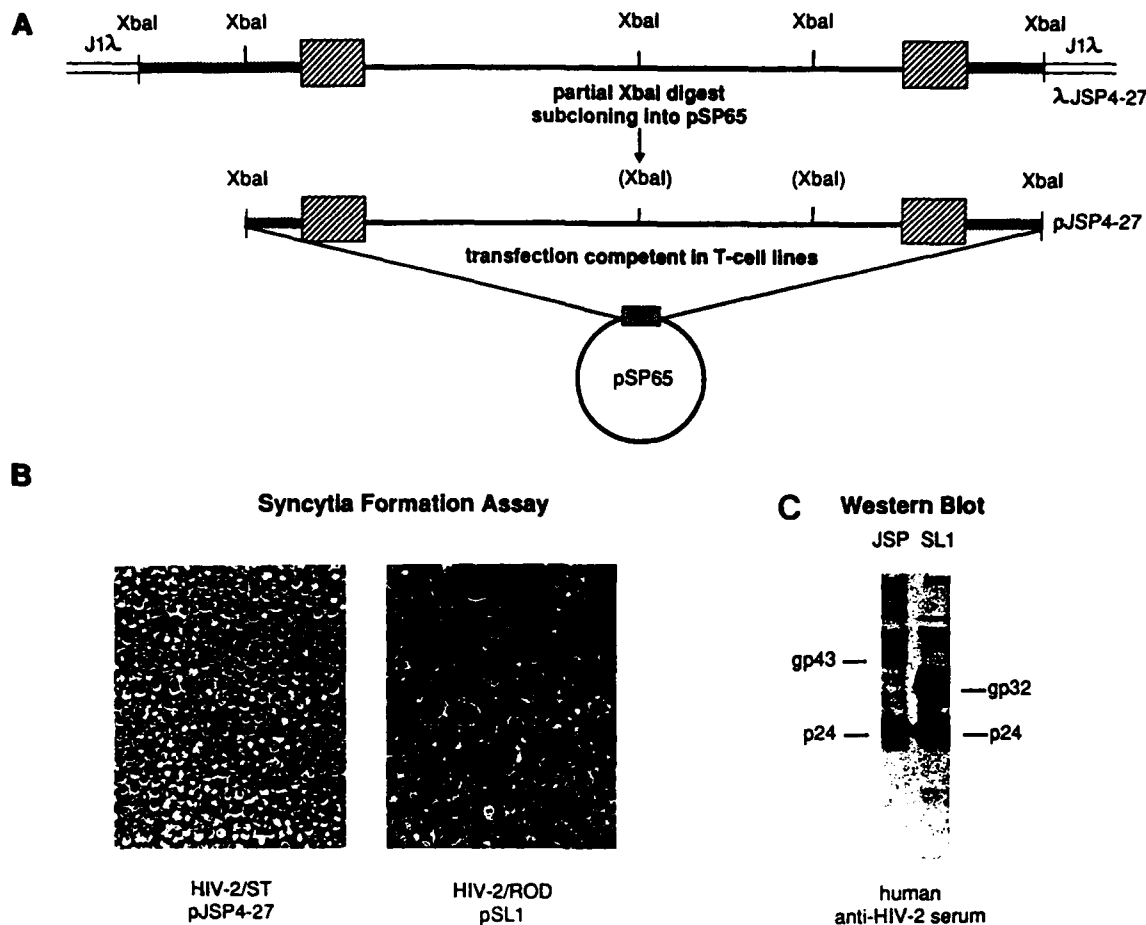
Figure 1. **A: Construction of a replication-competent HIV-2/ST plasmid clone** . The construction scheme used to generate a transfection-competent plasmid subclone of HIV-2/ST is outlined. λJSP4-27 (34) was partially cleaved with Xba I to remove phage arms (double lines) and flanking cellular sequences (hatched lines), and the resulting 14 Kb provirus-containing fragment was subsequently subcloned into pSP65. **B: Biological comparison of HIV-2/ST and HIV-2/ROD derived, genetically-pure viral strains.** CEMx174 cultures, productively-infected (>90%) with JSP4-27 (HIV-2/ST) and SL1 (HIV-2/ROD) respectively, were examined in syncytia formation assays and by Western blot analysis. The left side panels depict the complete lack of syncytium formation upon co-cultivation of JSP4-27-producing CEMx174 cells with uninfected CEM cells, while numerous and large syncytia are generated upon co-cultivation of the same uninfected CEM cells with SL1-infected CEMx174 cells (identical results were obtained with H9, SupT1 and CEMx174 cells). Syncytia formation was monitored 18 hours after cocultivation. The right side panel depicts the Western blot profile of cell-free virions derived from these same transfection-derived cultures and demonstrates differences in the size of their transmembrane envelope glycoproteins.

22

```
-->R                 .         .         .         .         .         .         .         .    100       .
AGTCGCTCTGCGGAGAGGCTGGCAGATTGAGCCCTGGGAGGTTCTCTCCAGCACTAGCAGGTAGAGCCTGGGTGTTCCCTGCTAGACTCTCACCAGTGCTTGGCCGGCACTGGGCAGACG
                                                    .R<---->U5.        .         200       .
GCTCCACGCTTGCTTGCTTAAAAGACCTCTTAATAAAGCTGCCAGTTAGAAGCAAGTTAAGTGTGTGCTCCCATCTCTCCTAGTCGCCGCCTGGTCATTCGGTGTTCATCTAAAGTAACA
                                        .    U5 <--|        .    PBS
AGACCCTGGTCTGTTAGGACCCTTTCTGCTTTGGGAAACCAAGGCAGGAAAATCCCTAGCAGGTTGGCGCCCGAACAGGGACTTGAAGAAGACTGAGAAGCCTTGGAACACGGCTGAGTG
                         400
AAGGCAGTAAGGGCGGCAGGAACAAACCACGACGGAGTGCTCCTAGAAAAGCGCAGGCCGAGGTACCAAGGGCGGCGTGTGGAGCGGGAGTGAAAGAGGCCTCCGGGTGAAGGTAAGTGC
          500                  .         .         .         .         .         .         600
CTACACCAAATACAGTAGCCAGAAGGGCTTGTTATCCTACCTTTAGACGGGTAGAAGATTGTGGGAGATGGGCGCGAGAAACTCCGTCTTGAGAGGGAAAAAAGCAGACGAATTAGAAAA
                                                   gag > MetGlyAlaArgAsnSerValLeuArgGlyLysLysAlaAspGluLeuGluLy
                    .         .         .         .         .         .   700       .         .
GATTAGGTTACGGCCCGGCGGAAAGAAAAAATATAGGCTAAAACATATTGTGTGGGCAGCGAATGAATTGGACAGATTCGGATTGGCAGAGAGCCTGTTGGAGTCAAAAGAGGGTTGCCA
sIleArgLeuArgProGlyGlyLysLysLysTyrArgLeuLysHisIleValTrpAlaAlaAsnGluLeuAspArgPheGlyLeuAlaGluSerLeuLeuGluSerLysGluGlyCysGl
                                        800
AAAAATTCTTACAGTTTTAGATCCATTAGTACCGACAGGGTCAGAAAATTTAAAAAGCCTTTTTAATACTGTCTGCGTCATTTGGTGTATACACGCAGAAGAGAAAGCCGAAAGATACTGA
nLysIleLeuThrValLeuAspProLeuValProThrGlySerGluAsnLeuLysSerLeuPheAsnThrValCysValIleTrpCysIleHisAlaGluGluLysAlaLysAspThrGl
                                        900
AGAAGCAAAACAAAAGGTACAGAGACATCTAGTGGCAGAAACAAAAACTACAGAAAAAATGCCAAGTACAAGTAGACCAACAGCACCACCTAGCGGGAACGGAGGAAACTTCCCCGTACA
uGluAlaLysGlnLysValGlnArgHisLeuValAlaGluThrLysThrThrGluLysMetProSerThrSerArgProThrAlaProProSerGlyAsnGlyGlyAsnPheProValGl
                    .         .         .         .   1000      .         .         .         .
ACAAGTGGCCGGCAACTATACCCATGTGCCACTAAGTCCCCGAACCCTAAATGCTTGGGTAAAACTAGTAGAGGAAAAGAAGTTCGGGGCAGAAGTAGTGCCAGGATTTCAGGCACTCTC
nGlnValAlaGlyAsnTyrThrHisValProLeuSerProArgThrLeuAsnAlaTrpValLysLeuValGluGluLysLysPheGlyAlaGluValValProGlyPheGlnAlaLeuSe
          1100                 .         .         .         .         .         .         1200
AGAAGGCTGCACGCCCTATGATATTAATCAAATGCTTAATTGTGTGGGCGACCATCAAGCAGCTATGCAAATAATCAGGGAAATTATTAATGAAGAAGCAGCAGATTGGGACGCACAACA
rGluGlyCysThrProTyrAspIleAsnGlnMetLeuAsnCysValGlyAspHisGlnAlaAlaMetGlnIleIleArgGluIleIleAsnGluGluAlaAlaAspTrpAspAlaGlnHi
                                                  .         .         1300      .         .
CCCAATACCAGGCCCCTTACCAGCGGGGCAGCTCAGGGAGCCAAGGGGATCTGACATAGCAGGGACAACAAGCACAGTAGAAGAGCAGATCCAGTGGATGTTTAGGCCACAAAATCCTGT
sProIleProGlyProLeuProAlaGlyGlnLeuArgGluProArgGlySerAspIleAlaGlyThrThrSerThrValGluGluGlnIleGlnTrpMetPheArgProGlnAsnProVa
                                        1400
ACCAGTAGGAAGCATCTATAGAAGATGGATCCAGATAGGGCTACAGAAGTGCGTCAGGATGTACAACCCAACCAACATCCTAGACATAAAACAGGGACCAAAGGAGCCATTCCAGAGTTA
lProValGlySerIleTyrArgArgTrpIleGlnIleGlyLeuGlnLysCysValArgMetTyrAsnProThrAsnIleLeuAspIleLysGlnGlyProLysGluProPheGlnSerTy
                                        1500
TGTAGATAGATTCTACAAGAGCTTGAGGGCAGAACAAACAGATCCAGCAGTAAAAAATTGGATGACCCAAACACTGCTAGTGCAGAATGCCAACCCAGACTGTAAGTTAGTACTAAAAGG
rValAspArgPheTyrLysSerLeuArgAlaGluGlnThrAspProAlaValLysAsnTrpMetThrGlnThrLeuLeuValGlnAsnAlaAsnProAspCysLysLeuValLeuLysGl
                                        1600
ACTAGGGATAAATCCTACCTTAGAAGAAATGCTAACCGCCTGTCAGGGGGTAGGTGGACCAGGCCAGAAAGCCAGATTAATGGCAGAAGCCTTAAAGGAGGCCATGGCACCAGCCCCCAT
yLeuGlyIleAsnProThrLeuGluGluMetLeuThrAlaCysGlnGlyValGlyGlyProGlyGlnLysAlaArgLeuMetAlaGluAlaLeuLysGluAlaMetAlaProAlaProIl
          1700                 .         .         .         .         .         .         1800
CCCATTTGCAGCAGCCCAACAGAGAAGGACAATTAAGTGCTGGAATTGCGGAAAGGAAGGGCACTCGGCAAGACAATGCCGAGCACCTAGAAGACAAGGCTGCTGGAAATGTGGCAAGGC
eProPheAlaAlaAlaAlaGlnGlnArgArgThrIleLysCysTrpAsnCysGlyLysGluGlyHisSerAlaArgGlnCysArgAlaProArgArgGlnGlyCysTrpLysCysGlyLysAl
                                        1900
AGGACACATCATGGCAAAATGCCCAGAAAGACAGGCGGGTTTTTTAGGGTTGGGCCCATGGGGAAAGAAGCCCCGCAATTTCCCTGTGGCCCAAATCCCGCAGGGGCTGACACCAACAGC
ArgThrHisHisGlyLysMetProArgLysThrGlyGlyPhePheArgValGlyProMetGlyLysGluAlaProGlnPheProCysGlyProAsnProAlaGlyAlaAspThrAsnSer
aGlyHisIleMetAlaLysCysProGluArgGlnAlaGlyPheLeuGlyLeuGlyProTrpGlyLysLysProArgAsnPheProValAlaGlnIleProGlnGlyLeuThrProThrAl
                                        2000
ACCCCCGATAGACCCAGTAGAGGACCTACTAGAGAAGTACATGCAGCAAGGGAAAAGGCAGAGAGAGCAGAGAGAGAGGCCATACAAAGAAGTGACAGAGGACTTCCTGCAGCTCGAGAA
ThrProAspArgProSerArgGlyProThrArgGluValHisAlaAlaArgGluLysAlaGluArgAlaGluArgGluAlaIleGlnArgSerAspArgGlyLeuProAlaAlaArgGlu
aProProIleAspProValGluAspLeuLeuGluLysTyrMetGlnGlnGlyLysArgGlnArgGluGlnArgGluArgProTyrLysGluValThrGluAspPheLeuGlnLeuGluLy
                                        2100
ACAAGAGACACCATGCAGAGAGACGACAGAGGACTTGCTGCACCTCAATTCTCTCTTTGGAAAAGACCAGTAGTCACAGCACATGTTGAGGGCCAGCCAGTAGAAGTTTTGCTAGACACA
ThrArgAspThrMetGlnArgAspAspArgGlyLeuAlaAlaProGlnPheSerLeuTrpLysArgProValValThrAlaHisValGluGlyGlnProValGluValLeuLeuAspThr
sGlnGluThrProCysArgGluThrThrGluAspLeuLeuGluHisLeuAsnSerLeuPheGlyLysAspGln***
                                        2200
GGGGCTGACGACTCAATAGTAGCAGGCGTAGAGTTAGGGAGCAATTATAGTCCAAAGATAGTAGGGGGAATAGGGGGATTCATAAATACCAAAGAATATAAAAATGTAGAAATAAGAGTA
GlyAlaAspAspSerIleValAlaGlyValGluLeuGlySerAsnTyrSerProLysIleValGlyGlyIleGlyGlyPheIleAsnThrLysGluTyrLysAsnValGluIleArgVal
          2300                 .         .         .         .         .         .         2400
TTAAATAAAAGAGTAAGAGCCACCATAATGACAGGTGATACCCCAATCAACATTTTTGGCAGAAACATTCTGACAGCCTTAGGCATGTCATTAAATCTACCAGTCGCCAAGATAGAACCA
LeuAsnLysArgValArgAlaThrIleMetThrGlyAspThrProIleAsnIlePheGlyArgAsnIleLeuThrAlaLeuGlyMetSerLeuAsnLeuProValAlaLysIleGluPro
                                        2500
ATAAAAATAATGCTGAAGCCAGGAAAGGATGGACCAAAACTGAGACAATGGCCCTTAACAAAAGAAAAAATAGAGGCACTAAAAGAGATCTGTGAGAAAATGGAAAGAGAGGGCCAGCTA
IleLysIleMetLeuLysProGlyLysAspGlyProLysLeuArgGlnTrpProLeuThrLysGluLysIleGluAlaLeuLysGluIleCysGluLysMetGluArgGluGlyGlnLeu
                                        2600
GAGGAGGCACCTCCAACTAATCCTTATAATACCCCCACATTTGCAATCAAGAAAAAGGACAAAAACAAATGGAGAATGCTAATAGATTTTAGAGAACTAAACAAGGTAACTCAAGACTTC
GluGluAlaProProThrAsnProTyrAsnThrProThrPheAlaIleLysLysLysAspLysAsnLysTrpArgMetLeuIleAspPheArgGluLeuAsnLysValThrGlnAspPhe
```

Figure 2: Complete nucleotide sequence of the HIV-2/ST proviral genome. 9,672 bp of nucleotide sequence are depicted along with the deduced amino acid sequences of the corresponding viral proteins. The sequence starts at the 5' cap site and ends with the 3' polyadenylation site of the viral RNA. The primer binding site (complementary to the $tRNA_{lys}$), the polypurin tract (PPT), as well as short inverted repeats which flank the LTRs are underlined. Core enhancer sequences (E), SP1 binding sites (Sp1), the tata box (TATAA) and the polyadenylation site signal (AATAAA) are shown. The U3/R and R/U5 boundaries, as well as the splice donor (SD) and splice acceptor (SA) sites, have been determined in analogy with HIV-2/ROD (23). The vpr open reading frame contains a premature in-frame TAA stop codon at position 5777, which is indicated by three asterisks. Sequence analysis was performed by the chemical degradation method according to Maxam and Gilbert (41) as well as by the dideoxynucleotide-chain termination method according to Sanger (48). The nucleotide sequence of JSP4-27 has been submitted to the AIDS Sequence Data Base, Los Alamos National Laboratories, as well as to Genebank and EMBL libraries.

23

FIGURE 2 (cont)

```
                                              2700                           .
ACAGAAATCCAGTTAGGAATTCCACACCCAGCAGGACTAGCCAAGAAGAAACGAATTACTGTCCTAGATGTAGGGGATGCTTACTTTTCCATACCACTACATGAGGATTTTAGACAGTAT
ThrGluIleGlnLeuGlyIleProHisProAlaGlyLeuAlaLysLysLysArgIleThrValLeuAspValGlyAspAlaTyrPheSerIleProLeuHisGluAspPheArgGlnTyr
                              2800
ACTGCATTTACTCTACCATCAATAAACAATGCTGAACCAGGAAAAAGATACATATATAAAGTCTCACCACAGGGATGGAAGGGATCACCAGCAATTTTTCAGTACACAATGAGGCAGGTC
ThrAlaPheThrLeuProSerIleAsnAsnAlaGluProGlyLysArgTyrIleTyrLysValSerProGlnGlyTrpLysGlySerProAlaIlePheGlnTyrThrMetArgGlnVal
                  2900                                                       2900                               3000
TTAGAACCATTCAGAAAAGCAAACCCGGATATCATTCTCATTCAGTACATGGATGATATCTTGATAGCCAGCGACAGGACAGATTTAGAACATGACAGAGTGGTTCTGCAGCTAAAGGAA
LeuGluProPheArgLysAlaAsnProAspIleIleLeuIleGlnTyrMetAspAspIleLeuIleAlaSerAspArgThrAspLeuGluHisAspArgValValLeuGlnLeuLysGlu
                                              3100
CTTCTAAATGGCCTGGGATTTTCCACCCCAGATGAGAAGTTCCAAAAAGACCCTCCATACCAATGGATGGGCTATGAACTGTGGCCAACTAAATGGAAGCTGCAAAGAATACAATTGCCC
LeuLeuAsnGlyLeuGlyPheSerThrProAspGluLysPheGlnLysAspProProTyrGlnTrpMetGlyTyrGluLeuTrpProThrLysTrpLysLeuGlnArgIleGlnLeuPro
                                        3200
CAAAAGGAAGTATGGACAGTCAATGACATCCAAAAACTGGTGGGTGTCTAAATTGGGCAGCACAAATCTACCCAGGGATAAAGACCAGAAACTTATGTAGGTTAATCAGAGGAAA·ATG
GlnLysGluValTrpThrValAsnAspIleGlnLysLeuValGlyValLeuAsnTrpAlaAlaGlnIleTyrProGlyIleLysThrArgAsnLeuCysArgLeuIleArgGlyLys·met
                        3300
ACACTCACAGAAGAGGTACAGTGGACAGAATTAGCAGAAGCGGAACTAGAAGAAAACAAAATCATCTTAAGCCAGGAACAAGAAGGATGCTATTACCAAGAGGAAAAGGAGCTAGAAGCA
ThrLeuThrGluGluValGlnTrpThrGluLeuAlaGluAlaGluLeuGluGluAsnLysIleIleLeuSerGlnGluGlnGluGlyCysTyrTyrGlnGluGluLysGluLeuGluAla
                    3400
ACAGTCCAAAAAGATCAAGACAATCAGTGGACATATAAGATACACCAGGGAGGAAAAATTCTAAAAGTAGGAAAAATATGCAAAGGTAAAAAATACCCACACCAACGGAGTCAGACTCCTA
ThrValGlnLysAspGlnAspAsnGlnTrpThrTyrLysIleHisGlnGlyGlyLysIleLeuLysValGlyLysTyrAlaLysValLysAsnThrHisThrAsnGlyValArgLeuLeu
        3500                                                                 3600
GCACAAGTAGTTCAAAAAAATAGGAAAAGAAGCACTAGTCATTTGGGGACGAATACCAAAATTTCACCTACCAGTAGAAAGAGATACCTGGGAACAGTGGTGGGATAACTACTGGCAAGTG
AlaGlnValValGlnLysAsnIleGlyLysGluAlaLeuValIleTrpGlyArgIleProLysPheHisLeuProValGluArgAspThrTrpGluGlnTrpTrpAspAsnTyrTrpGlnVal
ACATGGATCCCAGACTGGGACTTCATATCTACCCCGCCACTGGTCAGATTAGTATTTAACCTGGTGAAAGATCCCATACTAGGCGCAGAAACCTTCTACACAGATGGATCCTGCAATAAG
ThrTrpIleProAspTrpAspPheIleSerThrProProLeuValArgLeuValPheAsnLeuValLysAspProIleLeuGlyAlaGluThrPheTyrThrAspGlySerCysAsnLys
                                                        3800
CAATCAAGAGAAGGAAAAGCAGGATACATAACAGATAGAGGAAGAGACAAGGTGAGGCTATTAGAGCAAACCACCAATCAGCAAGCAGAATTAGAAGCCTTTGCGATGGCAGTAACAGAC
GlnSerArgGluGlyLysAlaGlyTyrIleThrAspArgGlyArgAspLysValArgLeuLeuGluGlnThrThrAsnGlnGlnAlaGluLeuGluAlaPheAlaMetAlaValThrAsp
                                3900
TCAGGTCCAAAGGCCAACATTATAGTAGACTCACAATATGTAATGGGAATAGTAGCAGGCCAACCAACAGAGTCAGAGAGTAAAATAGTAAATCAAATCATAGAAGAAATGATAAAAAAG
SerGlyProLysAlaAsnIleIleValAspSerGlnTyrValMetGlyIleValAlaGlyGlnProThrGluSerGluSerLysIleValAsnGlnIleIleGluGluMetIleLysLys
                            4000
GAAGCAATCTATGTTGCATGGGTCCCAGCCCATAAAGGCATAGGAGGAAATCAGGAGGTAGATCACTTAGTAAGTCAGGGCATCAGACAAGTATTATTCCTAGAGAAAATAGAACCCGCT
GluAlaIleTyrValAlaTrpValProAlaHisLysGlyIleGlyGlyAsnGlnGluValAspHisLeuValSerGlnGlyIleArgGlnValLeuPheLeuGluLysIleGluProAla
                4100                                                          4200
CAGGAGGAACATGAAAAATATCATAGCAATGTAAAAGAACTATCCCATAAATTTGGACTGCCCAAATTAGTGGCAAGACAAATAGTAAACACATGCACCCAATGTCAGCAGAAAGGGGAG
GlnGluGluHisGluLysTyrHisSerAsnValLysGluLeuSerHisLysPheGlyLeuProLysLeuValAlaArgGlnIleValAsnThrCysThrGlnCysGlnGlnLysGlyGlu
                                                4300
GCTATACATGGGCAAGTAAATGCAGAATTAGGCCACTTGGCCAAATGGACTGCACACACTTAGAAGGAAAAATCATTATAGTAGCAGTACATGTTGCAAGTGGATTTATAGAAGCAGAAGTC
AlaIleHisGlyGlnValAsnAlaGluLeuGlyThrTrpGlnMetAspCysThrHisLeuGluGlyLysIleIleIleValAlaValHisValAlaSerGlyPheIleGluAlaGluVal
                                    4400
ATCCCACAGGAATCAGGAAGGCAAACGGCACTCTTCCTACTAAAACTGGCCAGTAGGTGGCCAATAACACATTTGCACACAGACAATGGTGCCAACTTCACTTCACAGGAAGTAAAGATG
IleProGlnGluSerGlyArgGlnThrAlaLeuPheLeuLeuLysLeuAlaSerArgTrpProIleThrHisLeuHisThrAspAsnGlyAlaAsnPheThrSerGlnGluValLysMet
                                        4500
GTGGCATGGTGGATAGGTATAGAACAATCCTTCGGAGTACCTTACAATCCACAAAGCCAAGGAGTAGTGGAAGCAATGAATCACCACCTAAAAAATCAGATAAGCAGAATTAGAGAGCAG
ValAlaTrpTrpIleGlyIleGluGlnSerPheGlyValProTyrAsnProGlnSerGlnGlyValValGluAlaMetAsnHisHisLeuLysAsnGlnIleSerArgIleArgGluGln
                                            4600
GCAAACACAGTAGAAACAATAGTACTAATGGCAGTTCATTGCATGAATTTTAAAAGGAGGGGAGGAATAGGGGATATGACCCCAGCAGAAAGACTAATCAATATGGTCACTGCAGAACAG
AlaAsnThrValGluThrIleValLeuMetAlaValHisCysMetAsnPheLysArgArgGlyGlyIleGlyAspMetThrProAlaGluArgLeuIleAsnMetValThrAlaGluGln
            4700                                                             4800
GAAATACAATTCCTCCAAGCAAAAAAATTCAAAATTACAAAATTTTCGGGTCTATTTCAGAGAAGGCAGAGATCAGCTGTGGAAAGGACCTGGGGAACTACTGTGGAAGGGGGACGGAGCA
GluIleGlnPheLeuGlnAlaLysAsnSerLysLeuGlnAsnPheArgValTyrPheArgGluGlyArgAspGlnLeuTrpLysGlyProGlyGluLeuLeuTrpLysGlyAspGlyAla
                                                4900
GTCATAGTCAAGGTAGGGGCTGACATAAAAATAATACCAAGAAGGAAAGCTAAGATCATCAAAGACTATGGAGGAAGGCAAGAGATGGATAGCGGTTCCAACTTGGAGGGTGCCAGGGAG
ValIleValLysValGlyAlaAspIleLysIleIleProArgArgLysAlaLysIleIleLysAspTyrGlyGlyArgGlnGluMetAspSerGlySerAsnLeuGluGlyAlaArgGlu
                                                      vlf > MetGluGluGlyLysArgTrpIleAlaValProThrTrpArgValProGlyAr
                                                5000
GATGGAGAGGTGGCATAGCCTTATCAAGTATCTAAAATACAGAACAGGAGATCTAGAGAAGGTGTGCTATGTTCCCCACCATAAGGTGGGATGGGCGTGGTGGACTTGCAGCAGGGTAAT
AspGlyGluValAla···
gMetGluArgTrpHisSerLeuIleLysTyrLeuLysTyrArgThrGlyAspLeuGluLysValCysTyrValProHisHisLysValGlyTrpAlaTrpTrpThrCysSerArgValIl
                                        5100
ATTCCCATTAAAAAGGAGAAAGTCATCTGGAGATACAGGCATACTGGAACCTAACACCAGAAAAGGATGGCTCCTCCTATTCAGTAAGACTAACTTGGTATACAGAAAAATTCTGGAC
ePheProLeuLysGlyGluSerHisLeuIleIleGlnAlaTyrTrpAsnLeuThrProGluLysGlyTrpLeuSerSerTyrSerValArgLeuThrTrpTyrThrGluLysPheTrpTh
AGATGTTACCCCAGACTGTGCGGACTCCCTAATACATAGCACTTATTTCTCTTGCTTTACGGCAGGCGAAGTAAGAAGAGCCATCAGAGCGGAAAAGCTATTATCCTGCTGCAACTA·CC
rAspValThrProAspCysAlaAspSerLeuIleHisSerThrTyrPheSerCysPheThrAlaGlyGluValArgArgAlaIleArgGlyGluLysLeuLeuSerCysCysAsnTyr?r
            5300                                                            5400
CCAAGCCCATAAGTACCAGGTACCGTCACTCCAGTTTCTGGCCTTAGTGGTAGTGCAACAAAATGGCAGGCCCCAGAGAGACAATACCACCAGGAAACAGTGGCGAAGAAACTATCGGAG
oGlnAlaHisLysTyrGlnValProSerLeuGlnPheLeuAlaLeuValValValGlnGlnAsnGlyArgProGlnArgAspAsnThrThrArgLysGlnTrpArgArgAsnTyrArgAr
                                      vpx > MetAlaGlyProArgGluThrIleProProGlyAsnSerGlyGluGluThrIleGlyG
```

24

FIGURE 2 (cont)

```
                                                              .          .          5500        .          .
AGGCCTTCGAGTGGCTAGACAGGACGGTAGAAGCCATAAACAGAGAGGCAGTGAACCACCTGCCCCGAGAGCTTATTTTCCAGGTGTGGCAAAGGTCCTGGAGATACTGGCATGATGAAC
gGlyLeuArgValAlaArgGlnAspGlyArgSerHisLysGlnArgGlySerGluProProAlaProArgAlaTyrPheProGlyValAlaLysValLeuGluIleLeuAla***
luAlaPheGluTrpLeuAspArgThrValGluAlaIleAsnArgGluAlaValAsnHisLeuProArgGluLeuIlePheGlnValTrpGlnArgSerTrpArgTyrTrpHisAspGluG

                                              .          .          5600        .          .          .
AAGGAATGTCAATAAGTTACACAAAGTATAGATATTTGTGCCTAATGCAGAAAGCTATGTTCATACATTCTAAGAGAGGGTGCACTTGCCTGGGGGGAGGACATGGGCCGGGAGGATGGA
lnGlyMetSerIleSerTyrThrLysTyrArgTyrLeuCysLeuMetGlnLysAlaMetPheIleHisSerLysArgGlyCysThrCysLeuGlyGlyGlyHisGlyProGlyGlyTrpA
                                                        .          5700        .          .
GATCAGGACCTCCCCCTCCTCCCCCTCCAGGTCTAGTCTAATGACTGAAGCACCAACAGAGTCTCCCCCGGAGGATAGGACCCCACCGAGGGAGCCAGGGGATGAGTGGGTAATAGAAAC
rgSerGlyProProProProProProProGlyLeuVal***
                                          vpr >  MetThrGluAlaProThrGluSerProProGluAspArgThrProProArgGluProGlyAspGluTrpValIleGluTh
                                                      .          5800        .          .          .
CCTGAGAGAGATAAAATAAGAAGCTTTAAAGCACTTTGACCCTCGCTTGCTAATTACTCTTGGCAACTATATCTATGCTAGACATGGAGACACCCTTGAAGGCGCCAGAGGGCTCATTAG
rLeuArgGluIleLysEndGluAlaLeuLysHisPheAspProArgLeuLeuIleThrLeuGlyAsnTyrIleTyrAlaArgHisGlyAspThrLeuGluGlyAlaArgGlyLeuIleAr
                                                                       ***
                                                                                            tat >  MetGluThrProLeuLysAlaProGluGlySerLeuG
          5900        .          .          .          .          6000        .          .
GATCCTACAACGAGCCCTCCTCTTGCACTTCAGAGCAGGATGCGGCCGCTCAAGGATTGGTCAGCCCAGGGGACGAAATCCTTTATCAGCTATACCAACCCCTAGAGGCATGCGATAACA
gIleLeuGlnArgAlaLeuLeuLeuHisPheArgAlaGlyCysGlyArgSerArgIleGlyGlnProArgGlyArgAsnProLeuSerAlaIleProThrProArgGlyMetArg***
lySerTyrAsnGluProSerSerCysThrSerGluGlnAspAlaAlaAlaGlnGlyLeuValSerProGlyAspGluIleLeuTyrGlnLeuTyrGlnProLeuGluAlaCysAspAsnL
                                                                       .          6100        .          .
AATGTTACTGTAAAAAGTGCTGCTACCATTGCCAGATGTGTTTTTTAAACAAGGGGCTCGGGATATGGTATGAACGAAAGGGCAGAAGAAGAAGAACTCCGAAGAAAACTAAGGCTCATT
ysCysTyrCysLysLysCysCysTyrHisCysGlnMetCysPheLeuAsnLysGlyLeuGlyIleTrpTyrGluArgLysGlyArgArgArgArgThrProLysLysThrLysAlaHisS
                                                                                 rev >  MetAsnGluArgAlaGluGluGluGluLeuArgArgLysLeuArgLeuIle
          SD .          .          .          .          6200        .
CGTCTTCTGCATCAGACAAGTGAGTAAGATGTGTGGTAGGAATCAACTATTTGTTGCCAGCTTGCTAGCTAGTGCTTGCTTAATATATTGCGTCCAATATGTGACTGTTTTCTATGGCGT
erSerSerAlaSerAspLys
ArgLeuLeuHisGlnThrAsn
                              env >  MetCysGlyArgAsnGlnLeuPheValAlaSerLeuLeuAlaSerAlaCysLeuIleTyrCysValGlnTyrValThrValPheTyrGlyVa
                                                                       .          6300        .          .
GCCCGTGTGGAAAATGCATCCATTCCCCTCTTTTGTGCAACTAAAAATAGAGATACTTGGGGAACCATACAGTGCTTGCCAGACAATGATGACTATCAGGAAATAGCTTTAAATGTGAC
lProValTrpArgAsnAlaSerIleProLeuPheCysAlaThrLysAsnArgAspThrTrpGlyThrIleGlnCysLeuProAspAsnAspAspTyrGlnGluIleAlaLeuAsnValTh
                                        .          6400        .          .          .          .
AGAGGCCTTCGACGCATGGAATAATACAGTAACAGAACAAGCAGTAGAAGATGTCTGGAGTCTATTTGAGACATCAATAAAACCATGCGTCAAACTAACACCCTTATGTGTAGCAATGCG
rGluAlaPheAspAlaTrpAsnAsnThrValThrGluGlnAlaValGluAspValTrpSerLeuPheGluThrSerIleLysProCysValLysLeuThrProLeuCysValAlaMetAr
          6500        .          .          .          .          6600
TTGTAACAGCACAACTGCAAAAAAACAAACCTCCACACCAACAACCACCACAACAGCAAACACAACAATAGGAGAGAATTCTTCATGCATACGCACAGACAACTGCACAGGGTTGGGAGA
gCysAsnSerThrThrAlaLysAsnThrThrSerThrProThrThrThrThrThrAlaAsnThrThrIleGlyGluAsnSerSerCysIleArgThrAspAsnCysThrGlyLeuGlyGl
                                                            .          6700        .          .
AGAAGAGATGGTCGACTGTCAGTTCAATATGACAGGATTAGAGAGGGATAAGAAAAAACTATATAATGAAACATGGTACTCAAAAGATGTAGTCTGTGAATCAAATGACACCAAGAAAGA
uGluGluMetValAspCysGlnPheAsnMetThrGlyLeuGluArgAspLysLysLysLeuTyrAsnGluThrTrpTyrSerLysAspValValCysGluSerAsnAspThrLysLysGl
                                                  .          6800        .          .          .
GAAAACATGTTACATGAACCACTGCAACACATCAGTCATCACAGAGTCATGTGACAAGCACTATTGGGATACTATGAGGTTTAGATATTGTGCACCACCGGGTTTTGCCCTGCTAAGATG
uLysThrCysTyrMetAsnHisCysAsnThrSerValIleThrGluSerCysAspLysHisTyrTrpAspThrMetArgPheArgTyrCysAlaProProGlyPheAlaLeuLeuArgCy
                                        .          6900        .          .          .          .
CAATGATACCAATTATTCAGGCTTTGAGCCCAATTGTTCTAAGGTAGTAGCTGCTACATGTACAAGGATGATGGAAACGCAAACCTCCACTTGGTTTGGCTTTAATGGCACCAGGGCAGA
sAsnAspThrAsnTyrSerGlyPheGluProAsnCysSerLysValValAlaAlaThrCysThrArgMetMetGluThrGlnThrSerThrTrpPheGlyPheAsnGlyThrArgAlaGl
          7000        .          .          .          .          .
AAATAGAACATATATCTATTGGCATGGTAGGGATAATAGAACCATCATTAGCTTAAACAAGTTTTTATAATCTCACCGTACATTGTAAGAGGCCAGGAAACAAGACAGTTGTACCAATAAC
uAsnArgThrTyrIleTyrTrpHisGlyArgAspAsnArgThrIleIleSerLeuAsnLysPheTyrAsnLeuThrValHisCysLysArgProGlyAsnLysThrValValProIleTh
                    .          7100        .          .          .          .          7200
ACTCATGTCAGGGTTAGTGTTTCACTCCCAGCCAATCAATAGAAGACCCAGGCAAGCATGGTGCTGGTTCAAAGGCGAGTGGAAGGAAGCCATGAAGGAGGTGAAGCTAACCCTTGCAAA
rLeuMetSerGlyLeuValPheHisSerGlnProIleAsnArgArgProArgGlnAlaTrpCysTrpPheLysGlyGluTrpLysGluAlaMetLysGluValLysLeuThrLeuAlaLy
                                                            .          7300        .          .
ACATCCCAGGTATAAAGGAACCAACGACACAGAAAAAATTCGTTTTATAGCGCTAGGAGAACGCTCAGACCCAGAAGTGGCATACATGTGGACTAACTGCAGAGGAGAATTTCTCTACTG
sHisProArgTyrLysGlyThrAsnAspThrGluLysIleArgPheIleAlaLeuGlyGluArgSerAspProGluValAlaTyrMetTrpThrAsnCysArgGlyGluPheLeuTyrCy
                                        .          7400        .          .          .          .
CAATATGACTTGGTTCCTCAATTGGGTAGAAAACAGAACGAATCAGACACAGCACAATTATGTGCCATGCCATATAAAGCAAATAATTAATACCTGGCACAAGGTAGGGAAAAATGTATA
sAsnMetThrTrpPheLeuAsnTrpValGluAsnArgThrAsnGlnThrGlnHisAsnTyrValProCysHisIleLysGlnIleIleAsnThrTrpHisLysValGlyLysAsnValTy
                                        .          7500        .          .          .          .
TTTGCCTCCTAGGGAAGGACAGTTAACCTGCAACTCTACAGTGACCAGCATAATTGCTAACATTGACGGAGGAGAGAACCAGACAAATATTACCTTTAGTGCAGAGGTGGCAGAACTATA
rLeuProProArgGluGlyGlnLeuThrCysAsnSerThrValThrSerIleIleAlaAsnIleAspGlyGlyGluAsnGlnThrAsnIleThrPheSerAlaGluValAlaGluLeuTy
                                        .          7600        .          .          .          .
CCGATTAGAATTGGGGGATTATAAATTGATAGAAGTAACACCAATTGGCTTTGCACCTACACCAGTAAAAAGATACTCCTCTGCTCCAGTGAGGAATAAAAGAGGTGTATTCGTGCTAGG
rArgLeuGluLeuGlyAspTyrLysLeuIleGluValThrProIleGlyPheAlaProThrProValLysArgTyrSerSerAlaProValArgAsnLysArgGlyValPheValLeuGl
          7700        .          .          .          .          7800
GTTCTTAGGTTTTCTCACGACAGCAGGAGCTGCAATGGGCGCGGCCGTCCTTGACGCTGTCGGCTCAGTCTCGGACTTTATTGGCCGGGATAGTGCAGCAACAGCAACAGCTGTTGGACGT
yPheLeuGlyPheLeuThrThrAlaGlyAlaAlaMetGlyAlaAlaSerLeuThrLeuSerAlaGlnSerArgThrLeuLeuAlaGlyIleValGlnGlnGlnGlnGlnLeuLeuAspVa
                                        .          7900        .          .          .          .
GGTCAAGAGACAACAAGAAATGTTGCGACTGACCGTCTGGGGAACAAAAAATCTCCAGGCAAGAGTCACTGCTATCGAGAAATACTTAAAGGACCAGGCGCAACTAAATTCATGGGGATG
lValLysArgGlnGlnGluMetLeuArgLeuThrValTrpGlyThrLysAsnLeuGlnAlaArgValThrAlaIleGluLysTyrLeuLysAspGlnAlaGlnLeuAsnSerTrpGlyCy
                                        .          8000        .          .          .          .
TGCGTCTAGACAAGTCTGCCACACTACTGTACCATGGGTAAATGACACCTTAACGCCTGATTGGAACAACATGACATGGCAGGAATGGGAGCAACGAATCCGCAACCTAGAGGCAAATAT
sAlaSerArgGlnValCysHisThrThrValProTrpValAsnAspThrLeuThrProAspTrpAsnAsnMetThrTrpGlnGluTrpGluGlnArgIleArgAsnLeuGluAlaAsnIl
```

FIGURE 2 (cont)

```
                                                 .      8100      .           .           .           .           .           .
CAGTGAAAGTTTAGAACAGGCACAAATCCAGCCAAGAAAAGAACATGTATGAACTACAAAAATTAAATAGCTGGGATGTTTTTGGCAACTGGTTTGATTTAACCTCCTGGATCAAATATAT
eSerGluSerLeuGluGlnAlaGlnIleGlnGlnGluLysAsnMetTyrGluLeuGlnLysLeuAsnSerTrpAspValPheGlyAsnTrpPheAspLeuThrSerTrpIleLysTyrIl
                                                 .      8200      .           .           .           .           .           .
TCAGTATGGAGTTTATATAGTAGTAGGAATAATAGTTTTAAGAATAGTAATATATGTAGTACAAATGTTAAGTAGACTTAGAAAGGGCTATAGGCCTGTTTTCTCTTCCCCCCCCGCTTA
eGlnTyrGlyValTyrIleValValGlyIleIleIleValLeuArgIleValIleTyrValValGlnMetLeuSerArgLeuArgLysGlyTyrArgProValPheSerSerProProAlaTy
       SA       8300      .           .           .           .           .           .           .           .         8400
CTTCCAACAGATCCATATCCACAAGGACCGGGAACAGCCAGCCAGAGAAAGAAACAGAAGAAGACGTTGGAAACAGCGTTGGAGACAATTGGTGGCCCTGGCCGATAAGATATATACATTT
rPheGlnGlnIleHisIleHisLysAspArgGluGlnProAlaArgGluGluThrGluGluAspValGlyAsnSerValGlyAspAsnTrpTrpProTrpProIleArgTyrIleHisPh
          tat > SerIleSerThrArgThrGlyAsnSerGlnProGluLysLysGlnLysLysThrLeuGluThrAlaLeuGluThrIleGlyGlyProGlyArg***
          rev >   ProTyrProGlnGlyProGlyThrAlaSerGlnArgArgAsnArgArgArgArgTrpLysGlnArgTrpArgGlnLeuValAlaLeuAlaAspLysIleTyrThrPhe
                                                 .           .           .           .           .      8500      .           .
CCTGATCCGCCAGCTGATTCGCCTCTTGAACAGACTATACAACATCTGCAGGGACTTACTATCCAGGAGCTTCCAGACCCTCCAACTAATCTCCCAGAGTCTTCGGAGAGCATTGACAGC
eLeuIleArgGlnLeuIleArgLeuLeuAsnArgLeuTyrAsnIleCysArgAspLeuLeuSerArgSerPheGlnThrLeuGlnLeuIleSerGlnSerLeuArgArgAlaLeuThrAl
ProAspProProAlaAspSerProLeuGluGlnThrIleGlnHisLeuGlnGlyLeuThrIleGlnGluLeuProAspProProThrAsnLeuProGluSerSerGluSerIleAspSer
                                                 .           .           .      8600      .           .           .           .
AGTCAGAGACTGGCTGAGATTTAACACAGCCTACCTGCAATATGGGGGCGAGTGGATCCAAGAAGCGTTCCGAGCCTTCGCGAGGGCTACGGGAGAGACTCTTACAAACGCCTGGAGAGG
aValArgAspTrpLeuArgPheAsnThrAlaTyrLeuGlnTyrGlyGlyGluTrpIleGlnGluAlaPheArgAlaPheAlaArgAlaThrGlyGluThrLeuThrAsnAlaTrpArgGl
SerGlnArgLeuAlaGluIle***
                              nef > MetGlyAlaSerGlySerLysLysArgSerGluProSerArgGlyLeuArgGluArgLeuLeuGlnThrProGlyGluA
                                                 .           .      8700      .           .           .           .           .
CTTCTGGGGGACACTGGGACAAATTGGGAGGGGAATACTTGCAGTCCCAAGAAGGATCAGGCAGGGGGCAGAAATCGCCCTCCTGTGAGGGACGGCGGTATCAACAGGGAGATTTTATGA
yPheTrpGlyThrLeuGlyGlnIleGlyArgGlyIleLeuAlaValProArgArgIleArgGlnGlyAlaGluIleAlaLeuLeu***
laSerGlyGlyHisTrpAspLysLeuGlyGlyGluTyrLeuGlnSerGlnGluGlySerGlyArgGlyGlnLysSerProSerCysGluGlyArgArgTyrGlnGlnGlyAspPheMetA
                                                 .           .           .      8800      .           .           .           .
ATACCCCATGGAGAGCCCCAGCAGAAGGGGAGAAAGGCTCGTACAAGCAACAAAAATATGGATGATGTAGATTCAGATGATGATGACCTAGTAGGGGTCCCTGTCACACCAAGAGTACCAT
snThrProTrpArgAlaProAlaGluGlyGluLysGlySerTyrLysGlnGlnAsnMetAspAspValAspSerAspAspAspAspLeuValGlyValProValThrProArgValProL
                              .      8900      .           .      PPT  .     I-->U3.                    .           .      9000
TAAGAGAAATGACATATAGGTTGGCAAGAGATATGTCACATTTGATAAAAGAAAAGGGGGGACTGGAAGGGCTGTATTACAGTGATAGGAGACGTAGAGTCCTAGACATATACTTAGAAA
euArgGluMetThrTyrArgLeuAlaArgAspMetSerHisLeuIleLysGluLysGlyGlyLeuGluGlyLeuTyrTyrSerAspArgArgArgArgValLeuAspIleTyrLeuGluL
                                                 .           .           .           .      9100      .           .           .
AGGAAGAGGGAATAATTGGAGACTGGCAGAACTATACTCATGGACCAGGAGTAAGGTATCCAAAGTTCTTTGGGTGGTTATGGAAGCTAGTACCAGTAGATGTCCCACAAGAGGGAGATG
ysGluGluGlyIleIleGlyAspTrpGlnAsnTyrThrHisGlyProGlyValArgTyrProLysPhePheGlyTrpLeuTrpLysLeuValProValAspValProGlnGluGlyAspA
                                                 .           .           .           .      9200      .           .           .
ACAGTGAGACTCACTGCTTAGTGCATCCAGCACAAACAAGCAGGTTTGATGACCCGCATGGAGAAACATTAGTTTGGAGGTTTGACCCCACGCTAGCTTTTAGCTACGAGGCCTTTATTC
spSerGluThrHisCysLeuValHisProAlaGlnThrSerArgPheAspAspProHisGlyGluThrLeuValTrpArgPheAspProThrLeuAlaPheSerTyrGluAlaPheIleA
                                                 .           .           .      9300      .           .           .           .
GATACCCAGAGGAGTTTGGGTACAAGTCAGGCCTGCCAGAGGATGAATGGAAGGCAAGACTGAAAGCAAGAGGGATACCGTTTAGCTAAAAACAGGAACAGCTATACTTGGTCAGGGCAG
rgTyrProGluGluPheGlyTyrLysSerGlyLeuProGluAspGluTrpLysAlaArgLeuLysAlaArgGlyIleProPheSer***
                              .      E   9400      .     E     .        Spl ___Spl___  Spl           .           .           .
GAAGTAACTAACAGAAAACAGCTGAGACTGCAGGGACTTTCCAGAAGGGGCTGTTACCAGGGGAGGGACATGGGAGGAGCCGGTGGGGAACGCCCTCATACTTTCTGTATAAATGTACCC
                              U3<-+->R                                                                          .      9600
GCTACTCGCATTGTATTCAGTCGCTCTGCGGAGAGGCTGGCAGATTGAGCCCTGGGAGGTTCTCTCCAGCACTAGCAGGTAGAGCCTGGGTGTTCCCTGCTAGACTCTCACCAGTGCTTG

GCCGGCACTGGGCAGACGGCTCCACGCTTGCTTGCTTAAAAGACCTCTTAATAAAGCTGCCAGTTAGAAGCA
```

26

```
 •             • ••                    •                •                                                                      •    •   •    128
MCG.RNQLFVASLLASACLIYCVQYVTVFYGVPVWRNASIPLFCATKNRDTWGTIQCLPDNDDYQEIALEVIEAFDAWNNIVTEQAVEDVWSLFETSIRPCVKLTPLCVAMRCNSI...TAKNIIS.T   HIV-2/ST
-M...---LI-I-------V--T----------T-K--T-------R-----------------------T----------------I----N-----------------K-S--ESS-GN----KS       HIV-2/ROD
-S--.KI--L--F--T-------TK---------K-----------------------------------P-----------D-I--------N--------------T-N--........NAS-ESA       HIV-2/ISY
-K-SK---LI-IV---Y--H-K-F-------I-A-------------------------T------------------------------------------N--............-RNN-....       HIV-2/NIH_z
---.KSL-C---------Y-V--T---------------------------K-------T-------D---------------------------S---..---.N----..-                     HIV-2/GN
-GCLG---LI-I--L-VYG----T----------A----T----------------T------G--S-L------S----E-------I----Q------------S---IT----KSETDRWGL-K-S-    SIV/MAC
-GCLG---LI-L--V-VLE-C------------A-K--T-------------T---------S-L-I---------D-------:----N---------------I-----K-ETDRWGL-GWAG           SIV/SM


               •                     •                                    •   •  •                                         •        256
PTITTTANII.......IGENSS.CIRTDNCIGLGEEEHVDCQFNNIGLERDKKKLYNEIWYSKDVVCES.NDIKKEKTCYHNHCNISVITESCDKHYWDTHRFRYCAPPGFALLRCNDINISGFEP       HIV-2/ST
TS----TP-DQEQE....-S-DT.P-A-A---S-------TIN-------------Q------------TN-S-.NQTQ----------------AI---------Y--------------A-          HIV-2/ROD
VA--SPSGP........DMINDTDP--QLN--S--R--D--E---------L----Q-S----------D-S-.DR-R---------------------A----------V----------          HIV-2/ISY
TW--GR-.D--QNIT....-INDT-.HA-A-----K----I----S--------R-Q-T-A--------D.N-TS.SQSK----------------------A-------------------A-          HIV-2/NIH_z
.-GS--GHSE........-N-.T-.PSYS-------K--I-N---Y-----------Q--------------N-TKDGKNR-----------------AIK----------Y-----------           HIV-2/GN
T.I--A-P-SAPVSEKIDMHVNET-S--AQN-----EQ-Q-IS-K-T-----K---T-E--------T-L----QG-S-DN-SR---------Q---------I--------Y-----------M-         SIV/MAC
T.---AIT--ATPSVAENVINE-NP--KNNS-A--EQ-P-IG-K------N----E--------R-LI---QS.ANES-SK---H------Q---------AI---------Y-------S--L--A-       SIV/SM


  •                      •                   •                   •  •  •            •                            •          384
NCSKVVAATCTRMMETQTSTWFGFNGIRAENRIYIYWHGRDNRIIISLNKFYNLIVHCKRPGNKIVVPITLHSGLVFKS..QPI.NRRPRQANCWFKGEWKEAMKEVKLTLAKHPRYKGT.NDIEKIR     HIV-2/ST
---------S-----------------------------Y---$L----------I-KQ-M----H----HY---.-K------------R--D--Q---E--------R--.---RN-S             HIV-2/ROD
---------S---------P---L-----------------Y----IL-R--E------------RR---..-K-I-KK------R-----R---Q---Q--V--------.---N--N             HIV-2/ISY
--------------------------------K--------N-----M------------L---F---FK---..--VI-KK--------E-Q------Q--E-----------NRSR--N-K          HIV-2/NIH_z
K------S--------------------------------Y---SI----------------------.----.T----------KER--Q---Q-I--------.---KN--N                  HIV-2/GN
K-------VSS--------------------------------Y----MK-R--------L-V-I---------.--.-D--K------G-R--D-I----Q-IV-----T--.-N-D--N            SIV/MAC
------VSS---------------------KS---------Y----MR-R--E----L-V-I---------.-.--.-E--K------E-S--K-IQ---E--V-----T--.---R--N             SIV/SM


          •    •                               •                                                   •                      512
FIALGERSDPEVAYHWTNCRGEFLYCNNIWFLNWVENR.......TNQ.....TQHNYVPCHIKQIINTWHKVGKNVYLPPREGQLTCNSIVTSIIANID.GGENQINIIFSAEVAELYRLELGDYKL    HIV-2/ST
-A-P-KG------------------------------I-...........K-HR--A--------------R--------E-S-------------WQNN--------------------           HIV-2/ROD
-T-PEKD-------------------------K.....-G-...,-----------E-----------------------E-S-E------------VD-D-R------------------          HIV-2/ISY
-K-P-RG----T---------S----------.......-G-.....K-R--A--R-R-------R----L-------E------------A-D.---------A--------------            HIV-2/NIH_z
-TKP-RG-------------------------P--.....---A------------R-------------------------------VNS-----------------------------           HIV-2/GN
LT-P.GGG----TF------------K-N-------D-D.....V-T-RPKERHRR-------R--------------------D--------L-----WTDG---S--M---------------         SIV/MAC
LT-P.AGG----TF------------K-N-------D-DQKGGRWKQ-NRKEQQKK-----R------------------------D--------L--E--WINS-E----M----------           SIV/SM


extracellular domain <--|--> transmembrane domain                                                           •  •   640
IEVTPIGFAPTPVKRYSSAPV.RNKRGVFVLGFLGFLTTAGAAMGAASLTLSAQSRTLLAGIVQQQQQLLDVVKRQQEMLRLTVWGTKNLQARVTAIEKYLKDQAQLNSWGCASRQVCHTTVPWVNDI    HIV-2/ST
V-I---------KE------HG.-HT-------A---S----------V--------------------------------L--------------Q--R-------F-------------S          HIV-2/ROD
V-----------AE-------G.-H----L-----------------------------FR------------------------------------A---R-------F-------------        HIV-2/ISY
V-I---------S-------HQ.-HT-------A---S--------------------------------------------------------------------------F-----S------        HIV-2/NIH_z
------------RE------.------------A---S--------------------------------------------------------------------------F---------S          HIV-2/GN
V-I----L---D----TTGGTS-----------A---S--------T-----------------------------L--------T---------------A---F---------P-AS            SIV/MAC
--I----L--S-R--TTTGAS-----------A---S------V---------------------------------L-----------T----------------------F--------P-E-        SIV/SM


                                                                                                              768
LTPDWNNMIWQEWEQRIRNLEANISESLEQAQIQQEKNMYELQKLNSWDVFGNWFDLTSWIKYIQYGVYIVVGIIVLRIVIYVVQMLSRLRKGYRPVFSSPPAYFQQIHIHKDREQPAREETEEDVGN    HIV-2/ST
-A---D--------KQV-Y------K------------------------I----------V-------L-I-AV-A-----------------------G-I--------G---N------G-S        HIV-2/ROD
---E------------HK--F-------------------------------------------------M----VA-------------------.-L--G-I-------W---D----------       HIV-2/ISY
---------------KV-Y------Q--------------------T--L-F---VR-------V-------VA---------I--------------G-I---------Q---------S            HIV-2/NIH_z
-S-------------KQV-Y------Q-----------------------------------------V----A-I--L-----------------G-L-------T--G---N-G----DRD         HIV-2/GN
-------D------RKVDF--E--TAL--E-----------------------A--------I-V----V-L-------I----AK--Q----------S---IT-TQQ--PAL-T---GK-G-G-E      SIV/MAC
-V-N----------RQVDF-----TQL--E---------------I----------R------L-L-V-G----------A---Q----------V---P---GQ-P-TK-GE-G-G-D             SIV/SM


              •                                                                                          •   •  •          893
SVGDNWWPWPIRYIHFLIRQLIRLLNALYNICRDLLSRSFQTLQLISQSLRRALTAVRDWLRFNTAYLQYGGEWIQEAFRAFAR.ATGETLTNAWRGFWGTLGQIGRGILAVPRRIRQGAEIALL      HIV-2/ST
NG--RY-----A-------------T---S----------L-----Y-N-........-----LR--F----C-------Q-A--.--R----AG-C--L-RV-ER----------------          HIV-2/ROD
D--SRS-----E-------L-----T----S--------LYLI--PL.........----LKA-----C-------Q-L--.V-R----S-G-SL--A--R------------------            HIV-2/ISY
NG---RS-----A-------L-----TG-----------ISPI--P-F----Q-----I-----LKA-----C-------Q-L--.T-R---AG-G-DL-RA-QR--------------L---          HIV-2/NIH_z
DD-YDL#----------HL-T---TG--K-------TNSP-HR----....N----I-----LKA---------Q---K.T-R---AS--G-LCAAVQRV-------------------             HIV-2/GN
GG-NSS---Q-E------------TW-FSN--T----AY-I--P-L-R-SAT-RR--EV--TELT--------WSYFH--VQ-GW-S--.---AG---DL-E--RRG--W---I-------L-LT--       SIV/MAC
RG-SRS---Q-E------------TW-FSS---W-L--Y-I--PVL---STT-QR--EVI-IEI------WRYF---VQ-MWKL-.R---AS--GDI-E---RV------I-------L-LT--         SIV/SM
```

Figure 3. **Alignment of the deduced envelope amino acid sequences of seven different HIV-2 and SIV viruses.** The deduced amino acid sequences of the envelope open reading frames of HIV-2/ST (JSP4-27), HIV-2/ROD (23), HIV-2/ISY (18), HIV-2/NIH$_z$ (54), HIV-2/GH (24), SIV/MAC (8), and SIV/SM (27) have been aligned for maximum homology. The position of the presumed primary envelope precursor cleavage site, as identified for HIV-2/ROD (23), is shown. Filled circles indicate cysteine residues conserved among all seven envelope sequences, while asterisks indicate non-conserved cysteine residues. Potential N-linked glycosylation sites (NXS/T) in the HIV-2/ST envelope sequence are underlined. Crosses depict two in-frame stop codons which are present in the sequence of SIV/MAC and HIV-2/GH. Dots represent gaps which were introduced for optimal sequence alignment.

27

```
                              BamHI
                               CC
JSP4-27   AGAATTGGGGGATTATAAATTGATAGAAGTAACACCAATTGGCTTTGCACCTACACCAGTA
          Primer 1           LysLeuIleGluValThrProIleGlyPheAlaProThrProVal
                                                                            ST/B12      10/10
          ----------------------------------------C-----------T------       ST/24       10/10
                                                                Ser
          ----------------------------------------C-----------T------       ST/24.1C    1/9
                                                                Ser
          ----------------------------------------C-----------T--A---       ST/24.1C    8/9
                                                                SerIle
          ----------------------------------------C-----------T------       ST/24.2C    5/10
                                                                Ser
          ----------------------------------------C-----------T------       ST/24.2C    5/10
                                                                Ser
```

```
                                        ↓ cleavage site
JSP4-27   AAAAGATACTCCTCTGCTCCAGTGAGGAATAAAAGAGGTGTATTCGTGCTAGGGTTCTTA
          LysArgTyrSerSerAlaProValArgAsnLysArgGlyValPheValLeuGlyPheLeu
          -------------------------------------------G---------G--------     ST/B12      10/10
          -------------------------------------------G---------G--------     ST/24       10/10
          -------------------------------------------G---------G--------     ST/24.1C    1/9
          -------------------------------------------G---------G--------     ST/24.1C    8/9
          -------------------------------------------G---------G--------     ST/24.2C    5/10
          -------------------------------------------G---------G--------     ST/24.2C    5/10
```

```
                     •••        •••
JSP4-27   GGTTTTCTCACGACAGCAGGAGCTGCAATGGGCGCGGCGTCCTTGACGCTGTCGGCTCAG
          GlyPheLeuThrThrAlaGlyAlaAlaAlaMetGlyAlaAlaSerLeuThrLeuSerAlaGln
          ------------------------------------------------------------     ST/B12      10/10
          ------------------------------------------------------------     ST/24       10/10
          ------------------------------------------------------------     ST/24.1C    1/9
          ------------------------------------------------------------     ST/24.1C    8/9
          ------------------------------------------------------------     ST/24.2C    5/10
          ------------------------------------------------------------     ST/24.2C    5/10
```

```
JSP4-27   TCTCGGACTTTATTGGCCGGGATAGTGCAGCAACAGCAACAGCTGTTGGACGTGGTCAAG
          SerArgThrLeuLeuAlaGlyIleValGlnGlnGlnGlnGlnLeuLeuAspValValLys
          ------------------------------------------------------------     ST/B12      10/10
          ------------------------------------------------------------     ST/24       10/10
          --------C---------------------------------------------------     ST/24.1C    1/9
              Ser                                                          ST/24.1C    8/9
          ------------------------------------------------------------     ST/24.2C    5/10
          ---------------------------------------------------------G--     ST/24.2C    5/10
                                                                   Glu
```

```
                    PstI
                    G  G
JSP4-27   AGACAA...//...AACATGTATGAACTACAAAAATTAAATAGC
          ArgGln        Primer 2
          ------                                                           ST/B12      10/10
          ------                                                           ST/24       10/10
          ------                                                           ST/24.2C    5/10
          ------                                                           ST/24.2C    5/10
```

Figure 4. Generation of fusogenic and cytopathic variants of HIV-2/ST. A diagram illustrating the origin of biologically-distinct HIV-2/ST substrains is depicted. A total of five subcultures were established from the original HIV-2/ST isolate by limiting dilution cloning. All biologically-cloned subcultures, the transfection-derived JSP4-27 cell line, as well as the parental bulk culture produce non-fusogenic and non-cytopathic progeny virus (hatched boxes). The two fusogenic and cytopathic variants ST/24.1C and ST/24.2C, which were generated by repeated cell-free passage of ST/24 supernatants to uninfected SupT1 cells, are depicted in black boxes. Also indicated is the size of the envelope transmembrane glycoprotein (TM) for virions derived from each culture.

Figure 5. Nucleotide sequence variation in the fusion domain of cytopathic and noncytopathic HIV-2/ST strains. 230 bps of PCR amplified envelope sequence are compared between JSP4-27 and four other HIV-2/ST strains. The boundaries of the amplified fragments are indicated. The sequence of both primers used to amplify the genomic DNA of ST/B12, ST/24, ST/24.1C and ST/24.2C is underlined, and the bp changes which were introduced to generate Bam HI and Pst I cloning sites are indicated. Nucleotide substitutions are shown with respect to the JSP4-27 sequence and amino acid sequence changes unique for the cytopathic and fusogenic ST/24 strains are bold-faced. Asterisks depict the Thr and Ala substitutions previously identified to distinguish JSP4-27 from most other cytopathic HIV-2 viruses. The number of M13 clones analyzed per HIV-2/ST strain is listed, with frequencies referring to the proportion of clones which have an identical sequence.

29

A

Cleavage Site

| extracellular domain | transmembrane domain | HIV-2 Envelope Gene |

M 1 2 3 4 5 6 7

← 544 bp →

Cleavage Site

BamHI | Psti | PCR Amplified Region

■ Fusion Domain

B

In Vitro
Cytopathicity

Proviral Sequences

ROD   KLVEITPIGFAPTKEKRYSSAHGRHTR GVFVLGFLGFLATAGSAMGAASLTVSAQSRTLLAGIVQQQQQLLDVVKRQ   +

JSP4-27 --I-V--------PV------PV-MK- ------------T---A--------L----------------------   -

PCR Sequences

M13
clones

B12   --I-V--------PV------PV-MK- ------------T---A--------L----------------------   10/10 -

ST/24   --I-V--------SV------PV-MK- ------------T---A--------L----------------------   10/10 -

ST/24.1C{ --I-V--------SV------PV-MK- ------------T---A--------L----------------------   1/9 } +
        --I-V--------SI------PV-MK- ------------T---A--------L-----s---------------   8/9 }

ST/24.2C{ --I-V--------SV------PV-MK- ------------T---A--------L----------------------   5/10 } +
        --I-V--------SV------PV-MK- ------------T---A--------L-----------------K---   5/10 }

Figure 6. Amino acid sequence variation in the fusion domain of cytopathic and non-cytopathic HIV-2/ST strains. The relative location of the PCR amplified envelope fragments is shown in the context of the entire HIV-2 envelope open reading frame. Amplification products are depicted in the left panel. Lanes 1 and 7, uninfected PBL DNA (negative control); lane 2, ST/B12; lane 3, ST/24; lane 4 ST/24.1C; lane 5, ST/24.2C, and lane 6, SupT1/LK001 (positive control). The deduced amino acid sequences of the amplified fusion regions are compared to the corresponding sequences of HIV-2/ROD and JSP4-27. The number of M13 clones analyzed per HIV-2/ST strain is listed, with frequencies referring to the proportion of clones which have an identical sequence.

30

	HIV-2/ST	HIV-2/ROD	HIV-2/ISY	HIV-2/NIH$_z$	HIV-2/GH	Envelope Amino Acid Sequence Divergence SIV/MAC$_{142}$	SIV/SM
HIV-2/ST		19%	17%	18%	16%	28%	28%
HIV-2/ROD	11%		20%	19%	18%	28%	28%
HIV-2/ISY	10%	11%		20%	19%	30%	29%
HIV-2/NIH$_z$	12%	12%	13%		19%	28%	28%
HIV-2/GH	11%	12%	12%	15%		29%	29%
SIV/MAC$_{142}$	23%	23%	24%	26%	24%		19%
SIV/SM	23%	22%	23%	23%	23%	15%	
Total Nucleotide Sequence Divergence							

^{a} The percent nucleotide sequence divergence between HIV-2/ST (JSP4-27), HIV-2/ROD (22), HIV-2/ISY (17), HIV-2/GH (Hasegawa et al., in press), SIV$_{MAC_{142}}$ (8), and SIV$_{SM}$ (25) is shown along with the percent amino acid sequence divergence of their envelope glycoproteins. Sequences were aligned pairwise, using the Microgenie computer software (Beckman).

Table 1: Nucleotide and amino acid sequence divergence among HIV-2 and SIV viruses.
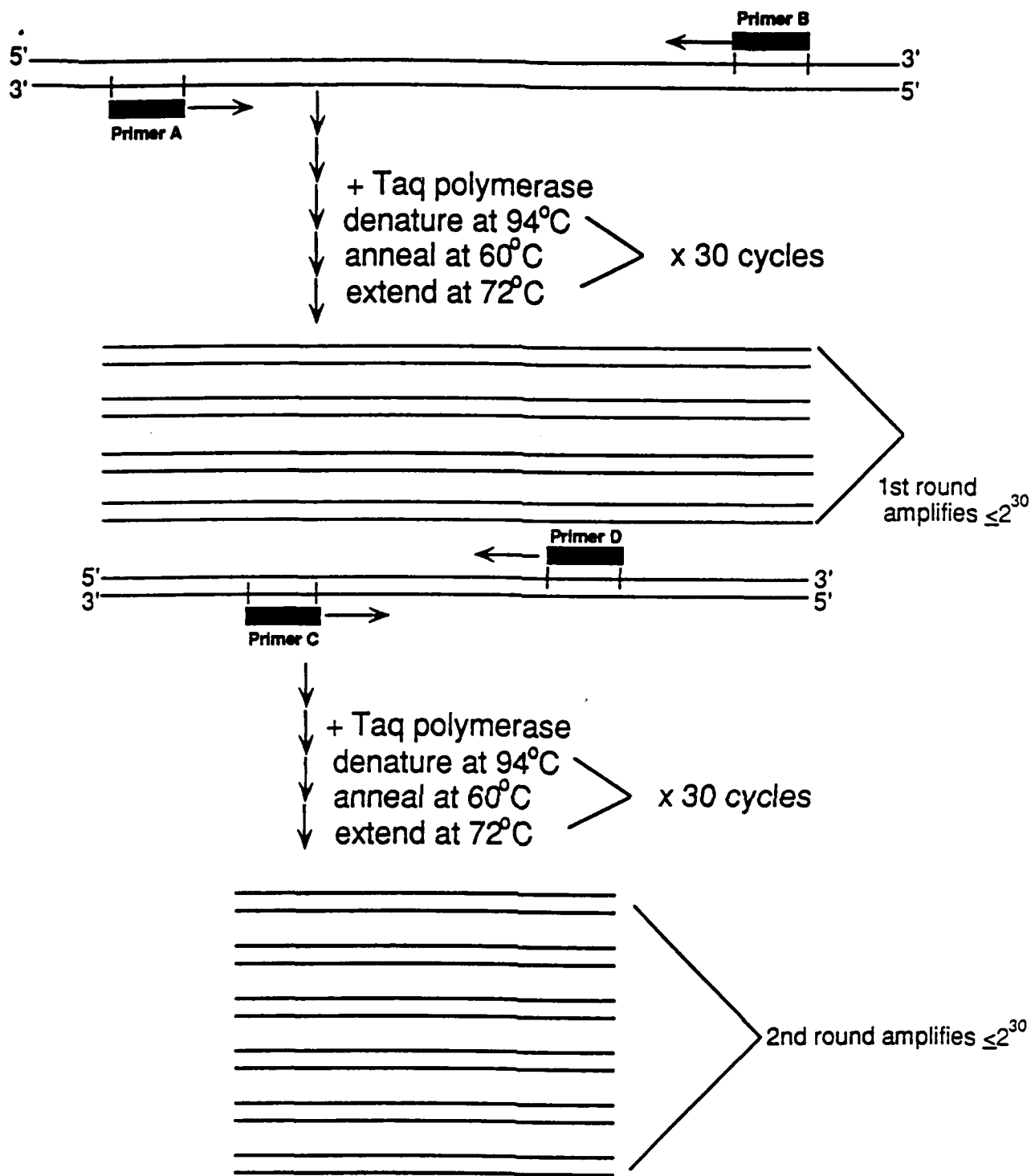
The % nucleotide sequence divergence between HIV-2/ST (JSP4-27), HIV-2/ROD, HIV-2/ISY, HIV-2/GH, SIV/MAC$_{142}$, and SIV/SM is shown along with the % amino acid sequence divergence of their envelope glycoproteins. Sequences were aligned pairwise using the Microgenie Computer Software (Beckman).

31

Open reading frame	ST/ROD		ST/ISY		ROD/ISY	
	% Homology					
	Nucleotide	Amino acid	Nucleotide	Amino acid	Nucleotide	Amino acid
LTR	91.5		91.4		91.8	
gag	91.1	92.0	90.7	89.1	90.3	89.5
pol	91.2	91.4	91.2	91.4	91.2	91.7
vif	91.5	88.4	92.4	91.2	92.2	85.6
vpx	89.3	85.7	92.6	90.2	88.4	87.5
vpr	89.9[a]	80.0[a]	89.8[a]	84.8[a]	93.3	89.5
tat	86.7	75.4	88.7	78.5	90.5	80.8
rev	82.0	85.0	ND[b]	ND[b]	ND[b]	ND[b]
env	85.5	81.4	86.0	83.0	84.8	80.4
nef	86.1	78.9	85.8	77.7	87.0	78.5
Overall % homology	89.5		89.9		89.3	

[a] An in-frame stop codon is present in the HIV-2/ST vpr open reading frame.
[b] Meaningful comparison of the HIV-2/ISY rev gene (17) with the corresponding rev genes of HIV-2/ST (JSP4-27) and HIV-2/ROD (22) was not possible because of considerable length differences between their sequences. ND, Not done.

Table 2: Sequence homologies among virus-specific genes of three different HIV-2 proviruses. Nucleotide and amino acid sequence homologies among different viral genes of HIV-2/ST (JSP4-27), HIV-2/ROD, and HIV-2/ISY are shown in comparison. A meaningful comparison of the HIV-2/ISY rev gene to the corresponding rev genes of HIV-2/ST and HIV-2/ROD was not possible, because of considerable length differences between their sequences. An asterisk indicates the presence an in-frame stop codon in the HIV-2/ST vpr open reading frame.

Figure 7. "Nested" polymerase chain reaction primer amplification as an approach to increase sensitivity of retroviral genome detection.

33

```
                                                    G
        4172  TCATGTGCCCAATGTCAACAGAAAGGGGAAGCTATACATGGGCAAGTAAATGCAGAACTA  HIV-2
              ||||||  |||||||| || |||||  |||||||||||||||||| |||||| |||| |||
        4129  ACATGTGATAAATGTCATCAAAAAGGAGAAGCTATACATGGGCAGGTAAATTCAGACCTA  SIVmac
                                        pol A

        4232  GGCACTTGGCAAATGGACTGCACACATTTAGAAGGAAAGATCATTATAGTAGCAGTACAT  HIV-2
              || |||||||||||||||||| || || ||| |||| |||||| ||  | ||||| ||||||
        4189  GGGACTTGGCAAATGGATTGTACCCATCTAGAGGGAAAAATAGTCATAGTTGCAGTACAT  SIVmac
                      EcoRI
                     ┌─────┐
                     │A   C│
        4292  GTTGCAAGTGGATTTATAGAAGCAGAAGTCATCCCACAGGAATCAGGAAGACAAACAGCA  HIV-2
              || || ||||||||||||||||||||| || ||||| ||| |||||||||| ||||||
        4249  GTAGCTAGTGGATTCATAGAAGCAGAAGTAATTCCACAAGAAACAGGAAGACAGACAGCA  SIVmac
              pol C

        4352  CTCTTCCTATTGAAACTGGCAAGTAGGTGGCCAATAACACACTTGCATACAGATAATGGT  HIV-2
              || || || || ||| ||||||| || |||||| || |||||  |||| |||||||||||||
        4309  CTATTTCTGTTAAAATTGGCAAGCAGATGGCCTATTACACATCTGCACACAGATAATGGT  SIVmac

 555│337  4412  GCCAACTTCACTTCACAGGAGGTGAAGATGGTAGCATGGTGGATAGGTATAGAACAATCC  HIV-2
              || ||||||  ||||| || || || ||||||||| |||||||||| ||| ||||| ||  ||
        4369  GCTAACTTTGCTTCGCAAGAAGTAAAGATGGTTGCATGGTGGGCAGGGATAGAGCACACC  SIVmac

        4472  TTTGGAGTACCTTACAATCCACAGAGCCAAGGAGTAGTAGAAGCAATGAATCACCATCTA  HIV-2
              |||||| |||||| |||||||||||||||| || ||||||||| |||||||||||||||| ||
        4429  TTTGGGGTACCATACAATCCACAGAGTCAGGGAGTAGTGGAAGCAATGAATCACCACCTG  SIVmac

        4532  AAAAACCAAATAAGTACAATCAGAGAACAGGCAAATACAATAGAAACAATAGTACTAATG  HIV-2
              |||||| |||||||  || ||||||| ||||| ||||||| ||  ||||||| ||||||| ||||
        4489  AAAAATCAAATAGATAGAATCAGGGAACAAGCAAATTCAGTAGAAACCATAGTATTAATG  SIVmac
                      EcoRI
              G      ┌─────┐ C
        4592  GCAATTCATTGCATGAATTTAAAAGAAGGGGGGGGAATAGGGGATATGACTCCATCAGAA  HIV-2
              ||| |||||||||||||||||||||||||||||||  ||| |||||||||||||||||| |||||
        4549  GCAGTTCATTGCATGAATTTTAAAAGAAGGGGAGGAATAGGGGATATGACTCCAGCAGAA  SIVmac
                                        pol D

        4652  AGATTAATCAATATGATCACCACAGAACAAGAGATACAATTCCTCCAAGCCAAAAATTCA  HIV-2
              |||||||| || |||||||| |||||||||||| ||||||| ||| | ||| | ||||| |||
        4609  AGATTAATTAACATGATCACTACAGAACAAGAAATACAATTTCAACAATCAAAAAACTCA  SIVmac

        4712  AAATTAAAAGATTTTCGGGTCTATTTCAGAGAAGGCAGAGATCAGTTGTGGAAAGGACCT  HIV-2
              ||||| ||| ||||||||||||||||| ||||||||||||||||||| ||||||| |||||
        4669  AAATTTAAAAATTTTCGGGTCTATTACAGAGAAGGCAGAGATCAGCTGTGGAAGGGACCC  SIVmac
                                        pol B
```
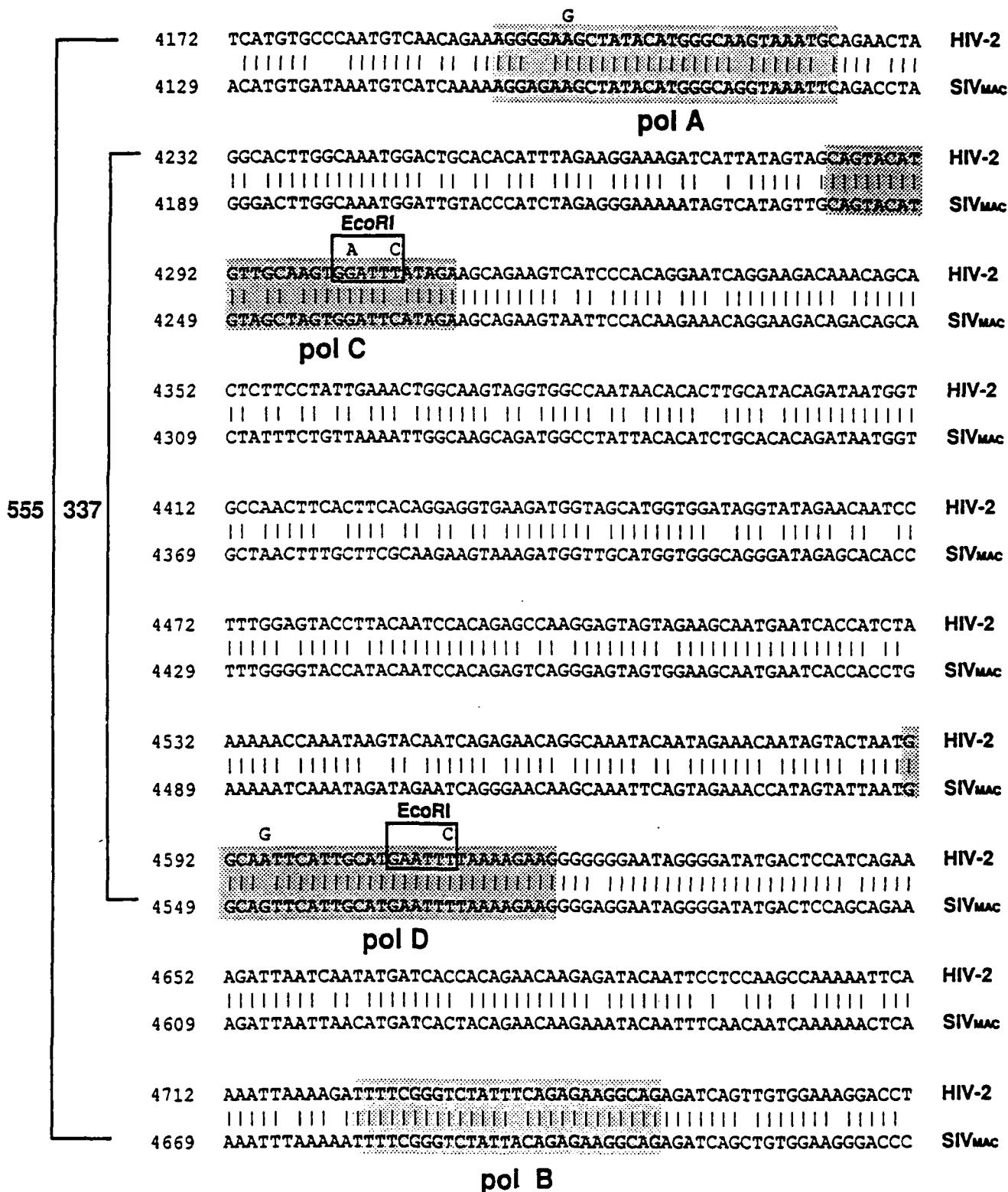
Figure 8. Nested HIV-2/SIV_mac pol primers with cloning enzyme sites. Outer (pol A and D) and inner (pol B and C) oligonucleotide primer sequences are shown in shaded areas and are derived from best match sequence of HIV-2 and SIV_mac. EcoRI sites were introduced within the inner primer sequences to facilitate cloning of the amplified product. The size of the outer and inner amplification products are 555bp and 337bp, respectively.

34

```
AGCAGAGAGTCATACCTAGGGAAACAGGAGAAAAGAAACGGCAAAGTTCTATTAAAAATACT   SIV AGMTYO
-----A--T--ACA----------------CT--A----C-C--C--G--G--GT--GC       SIV AGM385
-----A--C--ACA----------------CT--A----C-C--C--GC-G--GT--GC       SIV AGM384
G----------CAA----------------CT--A----C--------G--------T--GC    SIV AGM381

GAGTAGATGGCCTATAAACACAGTTACACACAGACAATGGGCCTAACTTTACCTCCCAAGA    SIV AGMTYO
C--C------C--T------CC-G--T------T------TA--T------AG--GC         SIV AGM385
C--C------C--C------C-G--T------T------TA-C------TAG--GC          SIV AGM384
A--C--G------C--TG--A-A----T--G--T------C--T------AG----C         SIV AGM381

AGTGGCAGCAATATGTTGGTGGGGAAAAATTGAACATACAACAGGTATACCATATAACCC     SIV AGMTYO
--A----T--C--C----------G--A----C----TTT--GG----C--C             SIV AGM385
--A----T--T--C------------A--G--C----TTT--GG----C--C             SIV AGM384
--A--------T----------G--A------------TTT--GG----C--C             SIV AGM381

CCAATCTCAAGGATCAATAGAAAGCATGAACAAGCAATTAAAAGAGATAATTGGGAAAAT     SIV AGMTYO
--AG--G--GT-G------TCT----------G----A--C--A--AC----             SIV AGM385
--AG--G--GT-G------TCT----------G----A--C--A--AC----             SIV AGM384
--GAG------GT-G------TC----------------T--A--AC-G----             SIV AGM381

AAGAGATGATTGCCAATATACAGAGGCAGCAGTACTGAT                           SIV AGMTYO
T--------GCAG--AGATTG--AA------CA-A--                             SIV AGM385
T--------GCAG--AGATTG--AA------CA-A--                             SIV AGM384
T--A---GCAG--AGATT----AA--------G--T--                            SIV AGM381
```

Figure 9. Nucleotide sequence comparison of pol gene segments of East (SIV_AGMTYO) and West (SIV_AGM385, 384, 381) African green monkey viruses. Nucleotide sequences were determined by the dideoxy chain termination method from three West African SIV_AGM virus strains (sabeus subspecies) and are shown compared to an East African strain (SIV_AGM-Tyo).

	HIV-1	HIV-2	SIV/SM	SIV/MAC	SIV/MND	SIV$_{AGM}$/GRI-1	SIV$_{AGM}$/VER-2	SIV$_{AGM}$/155	SIV$_{AGM}$/TYO	SIV$_{AGM}$SAB-4	SIV$_{AGM}$SAB-3	SIV$_{AGM}$SAB-2	SIV$_{AGM}$SAB-1
SIV$_{AGM}$/SAB-1	70.7	68.5	69.6	69.6	69.6	80.4	81.5	78.3	80.4	93.5	98.9	96.7	--
SIV$_{AGM}$/SAB-2	69.6	68.5	69.6	69.6	69.6	80.4	81.5	79.3	80.4	94.6	96.7	--	
SIV$_{AGM}$/SAB-3	70.7	69.6	70.7	70.7	69.6	81.5	80.4	79.3	79.3	93.5	--		
SIV$_{AGM}$/SAB-4	69.6	69.6	70.7	70.7	72.8	79.3	80.4	78.3	79.3	--			
SIV$_{AGM}$/TYO	62.0	60.9	62.0	62.0	67.4	80.4	93.5	92.4	--				
SIV$_{AGM}$/155	62.0	62.0	65.2	65.2	67.4	79.3	94.8	--					
SIV$_{AGM}$/VER-2	64.1	63.0	66.3	67.4	69.6	81.5	--						
SIV$_{AGM}$/GRI-1	69.6	62.0	63.0	63.0	68.5	--							
SIV/MND	64.5	62.0	64.1	64.1	--								
SIV/MAC	63.0	90.2	95.7	--									
SIV/SM	65.2	92.4	--										
HIV-2	62.0	--											
HIV-1	--												

Figure 10. Percent amino acid sequence identity in a 279bp PCR amplified pol region of primate lentiviruses. The sequences of the four West African SIV$_{AGM}$ viruses are indicated by SIV$_{AGM-sab1-4}$. The four SIV$_{AGM-sab}$ viruses differ from each other by 1-7% but from East African viruses by 19-22% and from isolates of HIV-1 and HIV-2 by 29-31%.
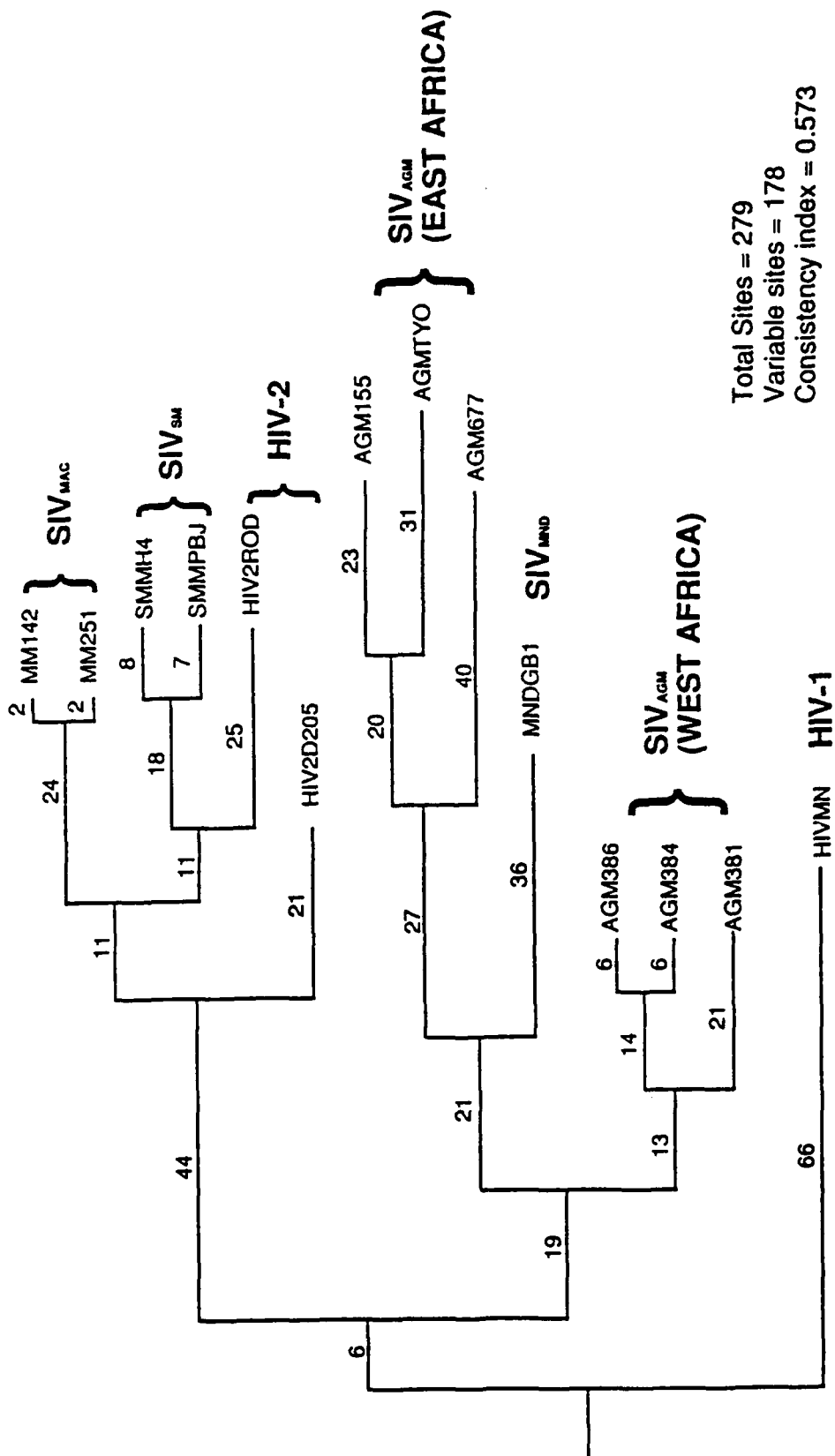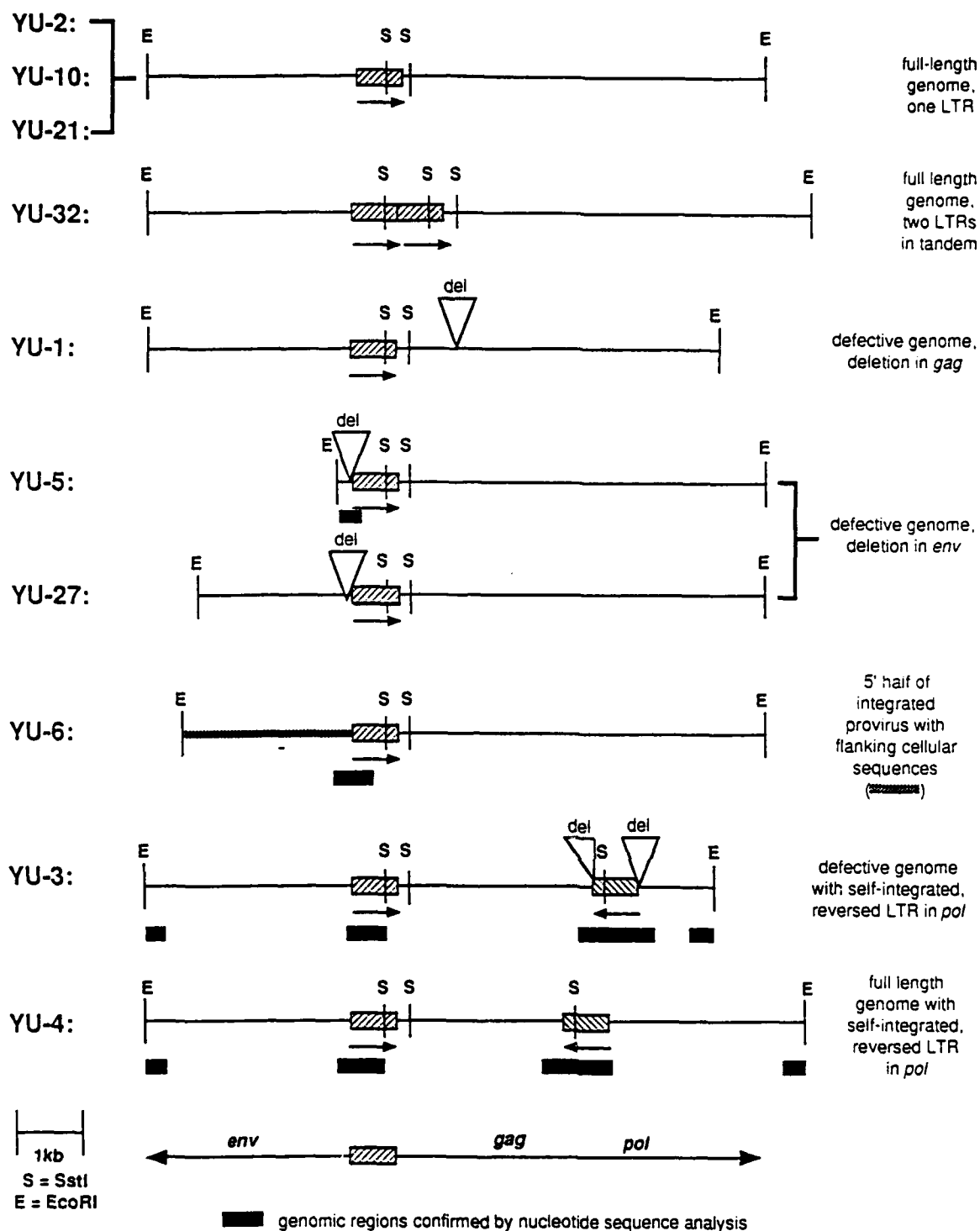
Figure 11. Phylogenetic tree analysis for human and simian immunodeficiency viruses.

Figure 12. Molecular proviral clones of HIV-1 derived from uncultured brain DNA of a patient with AIDS dementia. Ten HIV-1 proviral clones from a lambda phage genomic library of 8 x 10^6 recombinant phage were identified and mapped. Because the predominant viral forms in this uncultured brain DNA specimen were in unintegrated circular form, a single cutter, EcoRI, was selected as the cloning enzyme. Four of the 10 clones were full-length with either one or two LTRs (YU-2, 10, 21, 32), one clone was partial but integrated into genomic DNA (YU-6), and the remaining 6 clones were defective by virtue of deletions or rearrangements (YU-1, 3, 4, 5, 27).

38

DISTRIBUTION LIST

4 copies
Director
Walter Reed Army Institute of Research
ATTN: SGRD-UWZ-C
Washington, DC 20307-5100

1 copy
Commander
US Army Medical Research and Development Command
ATTN: SGRD-RMI-S
Fort Detrick, Frederick, Maryland 21702-5012

2 copies
Defense Technical Information Center (DTIC)
ATTN: DTIC-DDAC
Cameron Station
Alexandria, VA 22304-6145

1 copy
Dean
School of Medicine
Uniformed Services University of the Health Sciences
4301 Jones Bridge Road
Bethesda, MD 20814-4799

1 copy
Commandant
Academy of Health Sciences, US Army
ATTN: AHS-CDM
Fort Sam Houston, TX 78234-6100